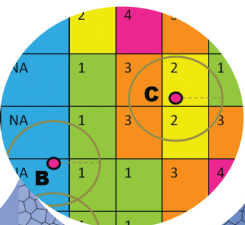
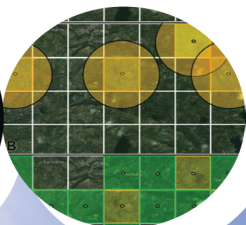


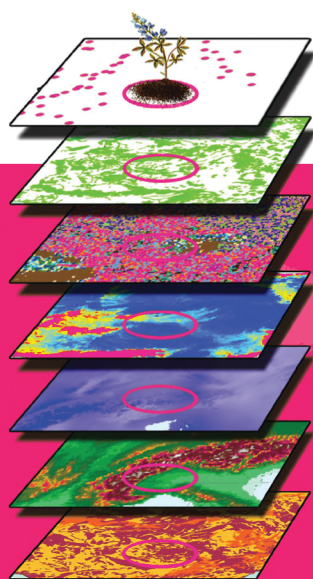
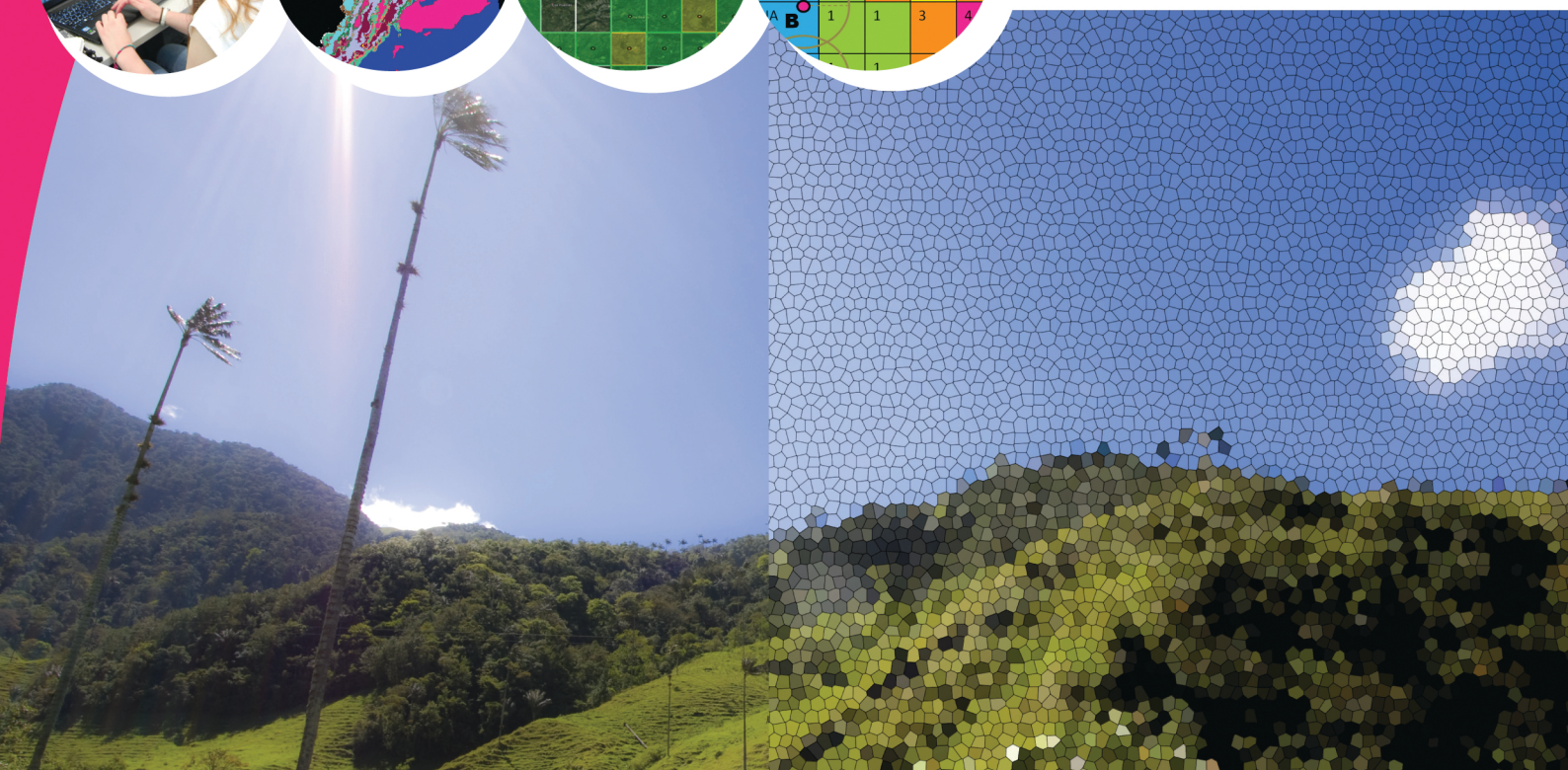


Tratado Internacional

SOBRE LOS RECURSOS FITOGENÉTICOS PARA LA ALIMENTACIÓN Y LA AGRICULTURA



Ferramentas



CAPFITOGEN

Programa para o Fortalecimento
das Capacidades em Programas
Nacionais de Recursos Genéticos
Vegetais da América Latina

Versão 1.2



Ferramentas

CAPFITOGEN

Programa para o Fortalecimento das Capacidades em Programas Nacionais de Recursos Genéticos Vegetais da América Latina

Versão 1.2

Autor das ferramentas

Mauricio Parra Quijano
Consultor
Tratado Internacional
sobre os Recursos Fitogenéticos
para a Alimentação e a Agricultura, TIRFAA
FAO

Autores do manual acompanhante

Mauricio Parra Quijano
Elena Torres Lamas, Universidad Politécnica de Madrid (España)
José María Iriondo Alegría, Universidad Rey Juan Carlos (España)
Francisco López, TIRFAA, FAO

Tradução (Português)

Cristiane Vendruscolo
Mario Angelo Vidor, EPAGRI, Santa Catarina, (Brasil)

As denominações empregadas nesse produto informativo e a forma como aparecem apresentados os conteúdos, não implicam da parte da Organização das Nações Unidas para Agricultura e Alimentação (FAO), juízo algum sobre a condição jurídica ou nível de desenvolvimento dos países, territórios, zonas ou cidades, ou de suas autoridades, nem a respeito das delimitações de suas fronteiras ou limites.

A menção de empresas ou de produtos de algum fabricante em particular, estejam ou não patenteados, não implica que a FAO os aprove ou recomende, em relação a outros de natureza similar que não tenham sido mencionados.

As opiniões expressas nesta publicação são de responsabilidade do (s) autor (es) e não refletem necessariamente as opiniões ou políticas da FAO.

ISBN

© FAO 2014

A FAO incentiva o uso, reprodução e divulgação do material contido neste produto de informação. Salvo indicação em contrário, o material pode ser copiado, baixado e impresso para estudo, pesquisa e ensino, ou para uso em produtos e serviços não comerciais, desde que se indique a FAO como fonte e detentora dos direitos autorais e não implique o endosso pela FAO das opiniões, produtos ou serviços dos usuários.

Todos os pedidos de tradução e direitos de adaptação, bem como revenda e outros direitos de uso comercial, devem ser feitos através de [www.fao.org/ contact-us/licence-request](http://www.fao.org/contact-us/licence-request) ou endereçados a copyright@fao.org.

Os produtos de informação da FAO estão disponíveis no site [www.fao.org/ publications](http://www.fao.org/publications) e podem ser adquiridos através de publications-sales@fao.org.

Produtos eletrônicos

FAO declina qualquer responsabilidade por erros ou deficiências no banco de dados ou software, ou documentação que vem com estes, manutenção e melhoria do programa e por qualquer dano que possa resultar a partir deles. A FAO também não será responsável por atualizar os dados e não assume nenhuma responsabilidade por erros ou omissões nos dados fornecidos. No entanto, os usuários são solicitados a relatar quaisquer erros ou deficiências neste produto a FAO.

Índice

	Pág.
1. Programa para o Fortalecimento das Capacidades em Programas Nacionais de Recursos Genéticos Vegetais da América Latina	1
2. Ferramentas CAPFITOGEN: características e instalação	3
3. Ferramenta GEOQUAL	17
4. Ferramenta ELC mapas	27
5. Ferramenta ECOGEO	39
6. Ferramenta Representa	49
7. Ferramenta DIV mapas	63
8. Ferramenta ColNucleo	91
9. Ferramenta FIGS_R	103
10. Erros freqüentes	119
11. Créditos	123
12. Anexos	127

1. Programa para o Fortalecimento das Capacidades em Programas Nacionais de Recursos Genéticos Vegetais da América Latina

Sob a proteção do Tratado Internacional sobre os Recursos Fitogenéticos para a Alimentação e Agricultura (TIRFAA) e da Agencia Española de Cooperación Internacional para el Desarrollo (AECID), entre os anos 2008 e 2010 se desenvolveram dois work-shops para a implementação do TIRFAA no grupo dos países da América Latina e do Caribe (GRULAC) em Cartagena de Indias (Colômbia, julho-agosto de 2008) e La Antigua (Guatemala, agosto de 2010). O êxito desta atividade demonstrou a efetividade deste tipo de work-shops em contribuir com a implementação dos objetivos do TIRFAA dentro da comunidade de países GRULAC. A coordenação dos work-shops entre a Secretaria do TIRFAA, a AECID, o Ministerio de Medio Ambiente y Medio Rural y Marino e o Centro Nacional de Recursos Fitogenéticos del Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria CRF-INIA tem sido um fator decisivo na consecução dos objetivos, dado o firme compromisso da Espanha com o TIRFAA e os estreitos laços existentes entre os Programas Nacionais de Conservação de Recursos Genéticos Vegetais e Institutos Nacionais de Pesquisa Agrícola (INIAs) dos países do GRULAC e Espanha.

A boa experiência colhida nos work-shops prévios e a importância de impulsionar alguns objetivos do TIRFAA na América Latina e Caribe, em particular os constantes nos artigos 5, 6, 7, 8 e 13.2 c, representou um incentivo chave para dar continuidade a este tipo de atividades. Ao mesmo tempo se fez evidente a necessidade de dar maior profundidade e conteúdo técnico à continuidade dos work-shops, estabelecendo um programa de transmissão de tecnologia, onde os work-shops representam um elemento dentro de uma estratégia de ação mais ampla.

Tomando como base estes antecedentes e as necessidades da região foi implementado o Programa para o Fortalecimento das Capacidades em Programas Nacionais de Recursos Genéticos Vegetais da América Latina - CAPFITOGEN. Este programa está enfocado no desenvolvimento de tecnologias apropriadas para países com abundante agrobiodiversidade e limitados recursos econômicos. Sua função é desenvolver a tecnologia, transferi-la e dar formação adequada ao pessoal técnico dos países da América Latina que são parte do Tratado.

A boa acolhida das ferramentas e metodologias que foram desenvolvidas sob a proteção do Programa CAPFITOGEN em 2013, fez com que em alguns países-objetivo do programa, fossem organizados work-shops de caráter nacional, por iniciativa e busca de financiamento pelos próprios interessados. Ao mesmo tempo

há surgido interesse de outros países e regiões que não eram objeto inicial do Programa, os quais vêm demandando tanto ferramentas como atividades de transferência e treinamento.

Desta forma, o Programa CAPFITOGEN espera servir não só como gerador e transmissor de tecnologia apropriada, como também, um modelo de transmissão em si mesmo. Um de seus aspectos mais inovadores é o fato de envolver aqueles que desenvolveram as metodologias científicas, a desenvolver as ferramentas facilitadas baseadas em suas metodologias e a realizar diretamente as atividades de transferência e treinamento técnico. Este modelo garante aos beneficiários do Programa, acesso direto aos científicos e desenvolvedores para a resolução de suas dúvidas ou exposição de seus casos. Ao mesmo tempo, os científicos se subsidiam diretamente das experiências e problemáticas dos técnicos dos Programas Nacionais, o que influenciará em futuras pesquisas de maior aplicação e mais ajustadas às necessidades reais.

2. Ferramentas CAPFITOGEN:

características e instalação

2.1. Origem

As plantas cultivadas foram em seu momento plantas silvestres cuja ampla carga genética lhes permitia resistir e se adaptar aos desafios que um ambiente mutável lhes apresentava, como uma nova praga, uma doença, um herbívoro, uma seca, etc. O processo de domesticação transformou estas plantas silvestres nos grãos, legumes, frutas e verduras que hoje conhecemos; produtos à medida das necessidades e gostos dos seres humanos. Mas para conseguir estes produtos teve-se que proceder a um intenso processo de seleção. Em consequência, a base genética das espécies domesticadas é bastante reduzida se a comparamos com a de seus ancestrais silvestres.

A evolução das espécies cultivadas tem deixado produtos entre os quais também encontramos diferenças. Por exemplo, as variedades e híbridos modernos, produzidos a partir da década de 60, são materiais com um potencial produtivo muito alto mas de base genética demasiado estreita, se comparamos com as variedades do princípio do século XX.

Assim, a domesticação, um processo precisamente seletivo, tem deixado pelo caminho muitas configurações genéticas em prol de obter a “melhor” variedade. Nesse processo têm desaparecido muitos gens valiosos que poderiam contribuir com soluções ante futuros problemas com nossas produtivas, homogêneas e vulneráveis variedades modernas. Felizmente, nem tudo foi eliminado dos campos e agricultores de todo o mundo seguem conservando variedades herdadas de seus antepassados, provavelmente não comerciais, mas com valores muito importantes para sua cultura, hábitos alimentícios ou inclusive religiosos.

Conscientes da perda progressiva desse patrimônio genético, muitas nações vêm resgatando e conservando essas variedades e plantas silvestres aparentadas, depositando-as em coleções de germoplasma (sementes, tecidos, propágulos) desde os anos 50.

2.1.1 Conservar para conhecer e utilizar

As coleções de germoplasma se diferenciam essencialmente de uma coleção de museu em que o conservado é para ser utilizado. O principal usuário são os melhoradores genéticos de cultivos, os quais buscam aí fontes de caracteres de interesse para serem transmitidos a variedades modernas. No entanto, o uso eficiente e efetivo do germoplasma só é possível mediante o conhecimento que

possamos gerar do mesmo. A obtenção desse conhecimento se leva a cabo através dos processos de caracterização e avaliação, os quais requerem grandes esforços em termos econômicos e logísticos por parte das entidades responsáveis de conservar a agrobiodiversidade. A nível nacional, estas atividades ocorrem através dos Programas Nacionais, e para o caso da América Latina, estes Programas estão normalmente a cargo dos Institutos Nacionais de Pesquisa Agrária (INIAs).

2.1.2 Técnicas adequadas para cenários com escassos recursos

Os métodos para coletar, conservar e caracterizar a agrobiodiversidade com padrões científicos vêm, normalmente, de regiões e centros onde os recursos econômicos, a infra-estrutura ou a capacitação de pessoal não representam maiores limitações. Isto tem feito com que em países em desenvolvimento estas metodologias não se possam aplicar, ou se se aplicam, não se podem estender a todo o germoplasma conservado. Esta situação contrasta com o fato de que nos países em desenvolvimento é onde se concentra a maior riqueza genética agrícola vegetal.

Este cenário tem feito com que alguns grupos de pesquisa no mundo tenham dirigido seus esforços a explorar metodologias de menor custo e complexidade, que se adaptem muito melhor às condições de Programas Nacionais de países em desenvolvimento. As alternativas metodológicas incluem o aproveitamento da informação ambiental dos lugares de coleta (ecogeográfica), por exemplo, para fazer estimativas da variabilidade genética que contenha o germoplasma ou estimar com maior êxito a probabilidade de encontrar gens de interesse. Mesmo assim, se concebe o uso dos sistemas de informação geográfica como meio de obtenção e aproveitamento de tais dados ecogeográficos. Dado que a maior parte da informação ecogeográfica e os programas informáticos para realizar as análises são de acesso gratuito, o investimento se reduz a um computador com uma configuração de tipo comercial e capacitação de pessoal. Se trata, então, de métodos compatíveis com cenários de recursos limitados, condição que acontece frequentemente em Programas Nacionais de países em desenvolvimento. Depois de demonstrar sua efetividade em casos de estudo publicados em revistas científicas internacionais, esta nova geração de metodologias foi eleita para ser adaptada em forma de ferramentas simplificadas e práticas, e ser finalmente transferido a Programas Nacionais das regiões selecionadas.

2.2 Características

Os processos que englobam as aplicações da ecogeografia e os sistemas de informação geográfica (SIG) na conservação e uso eficiente dos recursos genéticos vegetais, tal e como têm sido publicados em diversos meios de divulgação científica, implicam a utilização de numerosos programas informatizados, alguns deles de difícil acesso ou não gratuitos, em sequências metodológicas complexas. Portanto, o desenvolvimento de ferramentas práticas

que permitam aplicar estes avanços por parte de pessoal técnico sem maior formação no manejo e programação de programas estatísticos e SIG, constitui o principal desafio.

A solução reside na utilização de um programa informatizado que reúna as seguintes características:

- a) Incluir todas as análises estatísticas requeridas.
- b) Incluir todas as funções dos SIG, tanto para o manejo de dados georreferenciados como para sua análise.
- c) Manejar bases de dados e produtos das análises estatísticas e dos SIG.
- d) Permitir guardar todos os resultados das análises em formatos informatizados convencionais.
- e) Ser de livre acesso e distribuição.

Na atualidade, a comunidade científica dispõe da linguagem informatizada de programação chamada “R” (R. Cran, 2012) que cumpre com as características necessárias para o desenvolvimento das ferramentas CAPFITOGEN. A linguagem “R” é muito potente para a análise, mas conta com o principal inconveniente de requerer conhecimentos técnicos específicos para poder programar seu funcionamento. O Programa CAPFITOGEN tem superado este inconveniente envolvendo as equipes de pesquisa responsáveis das metodologias originais no desenvolvimento das ferramentas, em particular, na programação de R. A forma em que se programa R busca integrar todas as funções e análises requeridas sob um mesmo programa, mas de maneira que funcione com os formatos de dados pré-definidos, através dos quais se possa introduzir os dados de qualquer usuário nas rotinas de R e se consigam resultados particulares.

O último desafio na popularização do uso destas ferramentas era a simplificação da forma em que se introduzem ordens e dados e se tem alcançado graças ao desenvolvimento de uma interface simples, baseada em programação java e html e com o uso de servidores virtuais Tomcat. Esta solução é oferecida pela aplicação RWUI (<http://sysbio.mrc-bsu.cam.ac.uk/Rwui>), que ao ser utilizada no desenvolvimento das ferramentas CAPFITOGEN, se adapta com pequenas modificações.

2.3 Instalação e execução

A versão 1.1 das ferramentas CAPFITOGEN requeria um complexo processo de instalação que incluía a instalação manual de R, o ambiente de execução de Java e do servidor virtual Tomcat, assim como a modificação de variáveis do ambiente Windows. O velho procedimento exigia do usuário fazer a instalação manual dos

programas e as modificações das variáveis de ambiente, além de levar em conta o tipo de sistema operativo (32 o 64 bit) da versão do Windows.

A partir da versão 1.2, o conjunto de ferramentas CAPFITOGEN vem com um instalador que se encarrega de realizar todos os passos necessários da instalação de programas e modificação de variáveis de ambiente. Além disso, este instalador se encarrega também de descompactar e instalar todos os pacotes de R necessários para que se leve a cabo as análises.

Deste modo, se conseguiu um pacote de ferramentas pronto para usar; os passos para fazer a instalação e execução das ferramentas CAPFITOGEN são os seguintes:

a) Se as ferramentas foram baixadas do web site disponível para isso (detalhes em <http://www.agrobiodiversidad.org/blog/?p=1039>), haverá um arquivo com extensão .alz (compactado) com o corpo principal das ferramentas e outros arquivos .alz com a informação ecogeográfica para cada país ou região que se vai considerar como área de trabalho. Extraia todos os arquivos ou pastas que contenham estes arquivos compactados e então execute o instalador das ferramentas, colocando a pasta com o nome do país ou região na pasta “rdatamaps” do conjunto de pastas e arquivos do corpo principal das ferramentas, o qual se encontra englobado na pasta “CAPFITOGEN”. Copie esta pasta preferivelmente numa unidade de memória externa dedicada exclusivamente para as ferramentas ou diretamente no disco duro do PC, que frequentemente corresponde à unidade C:\. Os sistemas de downloads da internet e instaladores das ferramentas podem mudar em versões posteriores.

b) Se teve acesso às ferramentas através de uma memória USB fornecida pelo Programa CAPFITOGEN, simplesmente conecte-a a uma das portas USB de seu PC. Ao acessar o conteúdo da memória USB se encontrará uma pasta chamada “CAPFITOGEN”. Não mova o conteúdo desta pasta. Esta pasta contém a estrutura de pastas e arquivos (Fig. 1) necessária para o funcionamento das ferramentas.

c) No grupo de arquivos se encontrará um só arquivo executável (com extensão .exe), normalmente denominado “CAPFITOGEN.exe”. Clique duas vezes sobre este arquivo e imediatamente se abrirá uma janela como a que se mostra na Fig. 2.

Formatos	14/07/2013 07:12 ...
InstalarPrevia	14/07/2013 05:28 ...
Manuales	14/07/2013 07:13 ...
MapasApoyo	14/07/2013 07:24 ...
Packages	30/07/2013 10:43 a...
Pasaporte	14/07/2013 07:03 ...
ADMlevel	21/01/2013 10:51 a...
bioclim	23/02/2013 04:14 a...
edaph	11/04/2013 11:47 a...
gadm2	23/01/2013 02:52 ...
geophys	23/02/2013 04:14 a...
glc2000	20/01/2013 01:37 ...
HerramientasCAPFITOGEN	14/03/2013 06:16 ...
ID0	26/02/2013 01:31 ...
CAPFITOGEN	13/08/2013 10:49 a...
lista.países	11/07/2013 08:09 a...
ListaPaísesMar	26/02/2013 01:33 a...
mapamarlr	01/03/2013 03:59 ...

Figura 1. Conteúdo típico da memória USB ou arquivo compactado com as ferramentas CAPFITOGEN, onde se destaca o arquivo CAPFITOGEN.exe



Figura 2. Janela de instalação das ferramentas CAPFITOGEN.

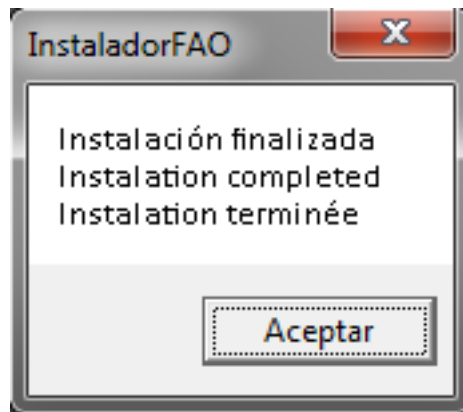


Figura 3. Janela que indica que a instalação foi finalizada.

d) Clique no botão “Instalar”. O instalador irá mostrando o progresso da instalação dos diferentes programas num texto de cor azul. Quando termine, aparecerá uma janela de cor cinza indicando a finalização do processo (Fig. 3). Clique no botão “Aceitar” desta janela. Não é necessário reiniciar o computador.

e) Imediatamente depois de terminar a instalação, na janela inicial, onde se encontrava o botão “Instalar”, aparecerá o botão “Executar”. Clique sobre este.

f) Nesse momento se abrirá uma janela de fundo preto na qual aparecerá uma série de código de cor branca (ver Fig. 4). Este código corresponde à programação para ativar o servidor virtual tomcat. Quando termina de executar as instruções aparece a seguinte linha de texto indicando os milissegundos usados para ativar o servidor virtual:

INFO: Server startup in xxxx ms

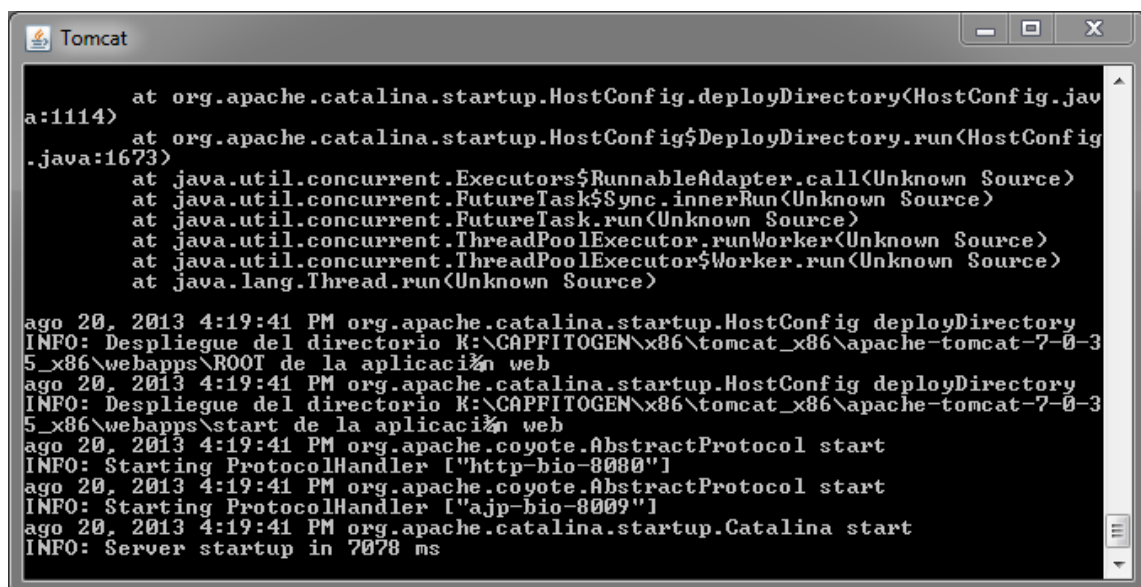


Figura 4. Janela de fundo preto onde se executa o servidor virtual. Esta janela não deve ser fechada até a finalização do uso das ferramentas.

g) Após vários segundos do início do processo anterior, se abrirá automaticamente o programa de navegação da internet pré-determinado (por exemplo: Internet Explorer, Mozilla Firefox ou Google Chrome) como mostra a Fig. 5. Na barra de links do navegador aparecerá a seguinte instrução:

<http://localhost:8080/start/>



Figura 5. O navegador se abre automaticamente mostrando o painel inicial.

Em algumas ocasiões o processo do ponto “f”, indicado anteriormente, inicia antes que o descrito no ponto “e” finalize. Isto faz com que o servidor virtual não esteja pronto quando é chamado pelo navegador, o qual gera um aviso de erro no navegador (ver Fig. 6). O adiantamento do processo “f” sobre o “e” se dá comumente na primeira vez que se executam as ferramentas ou quando a configuração do equipamento não está muito atualizada. Este problema se soluciona simplesmente fechando e abrindo o navegador e escrevendo na barra de links:

<http://localhost:8080/start/>

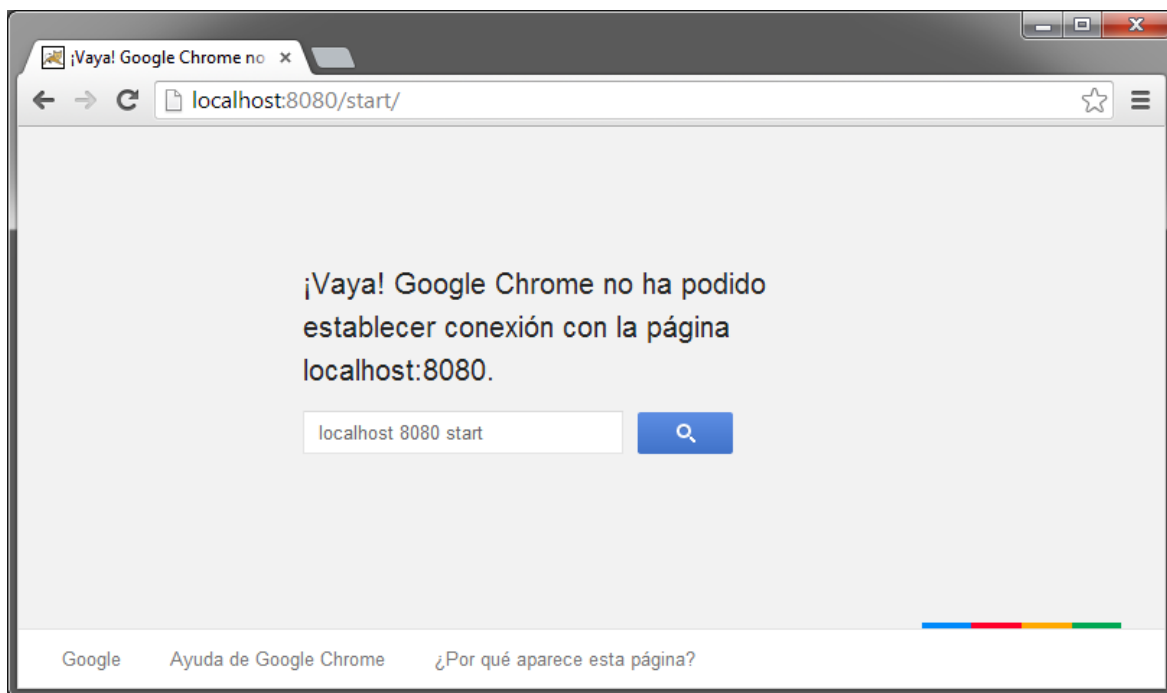


Figura 6. Exemplo de erro que se produz quando o navegador tenta abrir o link do servidor virtual sem que este tenha sido executado por completo.

h) Aparecerá no navegador um painel de início (Fig. 5), que mostra na parte esquerda a lista de ferramentas disponíveis e, na direita, os idiomas disponíveis para cada ferramenta. Ao clicar sobre os ícones de cor verde se abrirá no navegador o formulário da ferramenta e idioma selecionado. Na Fig. 7 mostra-se, como exemplo, o formulário da ferramenta GEOQUAL. Neste formulário aparece uma pequena barra de links que se explica na Fig. 8.

Para sair da aplicação, uma vez que se tenha terminado de utilizar as ferramentas, basta fechar o navegador, fechar a janela de fundo preto e fechar a janela da Fig. 2 e clicar no botão “sair”. É possível que depois de fechar todas estas janelas, o Windows pergunte se a aplicação foi corretamente instalada. Por favor, responda afirmativamente.

No momento em que se realiza a instalação, se adiciona um arquivo de identificação na unidade e caminho onde se encontram os arquivos e pastas das ferramentas CAPFITOGEN. Esse arquivo faz com que o computador reconheça que a instalação foi realizada. Desta maneira, cada vez que necessite usar de novo as ferramentas, pulse sobre o arquivo CAPFITOGEN.exe e se abrirá a janela de instalação, mas desta vez diretamente se apresentará o botão “Executar”. A partir deste passo se repete todo o processo desde o ponto “d” indicado anteriormente. Se por algum motivo a letra da unidade ou o caminho onde se encontra o arquivo executável muda, ou se elimina o arquivo de identificação acidentalmente, ao clicar duas vezes sobre CAPFITOGEN.exe aparecerá de novo o botão “Instalar”. É aconselhável, então, reinstalar. Se se

reinstalam novamente os programas R e Java não haverá problemas na execução das ferramentas.

Figura 7. Formulário da ferramenta GEOQUAL. 1. Links sobre aspectos legais. 2. Espaço que indica a ferramenta em uso. 3. Barra de links. 4. Descrição breve da ferramenta. 5. Link do manual de instruções. 6. Zona de introdução de parâmetros. 7. Botão para iniciar a análise.



Figura 8. Barra de links (corresponde à parte 3 da Fig.6). Os botões cumprem as seguintes funções: 1. Voltar ao formulário da ferramenta selecionada. 2. Ver manual de instruções das ferramentas CAPFITOGEN. 3. Informação de contato e apoio. 4. Retornar à página de seleção das ferramentas (página de aterrissagem).

i) Enquanto a análise está sendo executada, o Windows mostrará o ponteiro do mouse com o ícone de espera. Uma vez a análise concluída satisfatoriamente, a página que contém o formulário voltará a situar-se na parte superior, mostrando o cabeçalho com o logo do TIRFAA. Indo à parte inferior da página aparecerá uma

mensagem de “Proceso finalizado” (Fig.9), indicando que se deve buscar os resultados já guardados. Se se deseja mudar algum parâmetro e voltar a executar a análise, pode-se fazê-lo a partir do mesmo formulário, mudando o parâmetro em questão e clicando sobre o botão “Analisar”. Pode-se, também, deletar todo o conteúdo do formulário com o botão “Clear Page”.

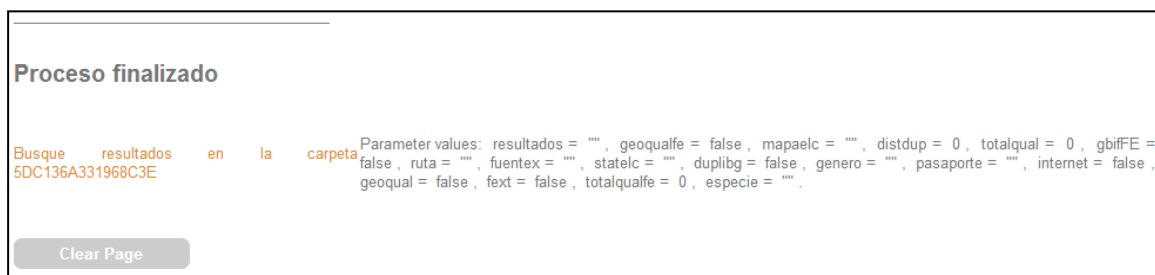


Figura 9. Vista da parte inferior da página do formulário com análise já concluída corretamente e com os resultados disponíveis.

Os resultados das análises devem ser buscados na pasta definida pelo usuário no parâmetro “resultados”. Os mapas, tanto em formato ráster (malhas de células com extensões .grd e .gri), - por exemplo, os mapas produzidos pela ferramenta ELC mapas -, como em formato vetorial (usualmente “shapefiles”), - por exemplo, os mapas de pontos, - pode-se abrir com o programa DIVA GIS, já que são plenamente compatíveis. As tabelas de resultado são feitas em formato texto separado por tabulações, as quais pode-se abrir a partir dos programas como Microsoft Excel ou planilhas de cálculo do Open Office.

j) É possível que quando se saia de uma ferramenta para o painel de início (usando o botão ➡ da Fig. 8) e voltando à mesma ferramenta ou simplesmente colocando de novo o link no navegador, apareça uma janela como a que mostra a Fig. 10 com instruções em inglês. Nesta janela se pergunta ao usuário se deseja voltar a executar uma ferramenta que previamente havia sido executada. Não haverá nenhum problema se se executa a ferramenta a partir desta janela. Para continuar é só clicar sobre o botão “Run new copy of” (executar nova cópia de) e o nome da ferramenta.



Figura 10. Janela que indica que se está tentando abrir uma ferramenta previamente aberta. Para poder continuar, simplesmente clicar sobre o botão “Run new copy of” e o nome da ferramenta.

k) Quando se produz um erro na execução da ferramenta, seja por uma deficiência na instalação, seja por um parâmetro introduzido erroneamente ou por defeitos nas tabelas que contêm a informação de passaporte ou caracterização, a página que contém o formulário apresenta no topo (da mesma maneira que quando o processo finaliza com sucesso), no entanto, o tempo que leva é consideravelmente inferior ao que leva o processo com êxito, sendo que na parte inferior da página, aparecerá uma mensagem de erro, como mostra a Fig. 10. É possível detectar a origem do erro através da mensagem de erro. Assim, quando aparecer esta mensagem, gerada diretamente pelo programa R, deve-se comparar seu conteúdo com a lista de mensagens do capítulo 10 “Erros freqüentes”.

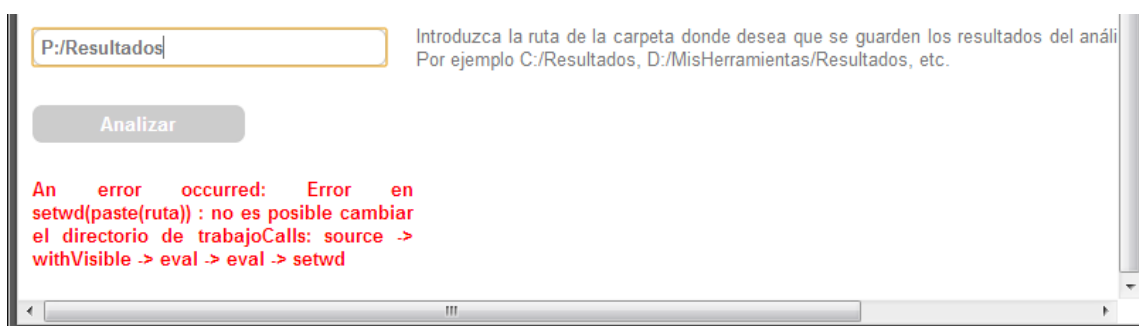


Figura 11. Vista da parte inferior da página do formulário quando se produziu algum tipo de erro.

l) Pelas configurações de alguns computadores pessoais, é possível que durante a instalação não se executem bem todos os pacotes que R requer para o alcance

das análises previstas. Isto produzirá um erro recorrente ao executar as ferramentas, o qual aparecerá em cor vermelha como o da Fig. 11 mas com o seguinte texto:

An error occurred: Error en library(nombre del paquete) : there is no package called 'nombre del paquete'
Calls: source -> withVisible -> eval -> eval -> library

Este erro, que aparece detalhado no capítulo 10 (Erros frequentes), pode se repetir para todos os pacotes necessários, tornando o processo de correção proposto no capítulo 10 extenso demais. Neste caso, se sugere que o usuário elimine a pasta “library” que encontrará no caminho C:\rwin, copie a pasta “library” que encontrará no caminho CAPFITOGEN\Error e finalmente cole-a no caminho C:\rwin. Esta ação eliminará por completo os problemas de ausências de pacotes necessários para R.

m) Em algumas ocasiões, devido a um processo deficiente de instalação ou a uma incorreta localização da pasta “CAPFITOGEN”, é possível que a interface não possa ativar R e enviar a informação necessária para realizar as análises. Este problema se fará evidente quando depois de clicar no botão “analisar” e haver preenchido corretamente os parâmetros, o tempo de execução seja extremamente curto e não apareçam mensagens de erro na cor vermelha, apenas a mensagem de “Processo finalizado”, sendo que na pasta onde se deveriam guardar os resultados, não apareça nenhum arquivo novo. Neste caso, assegure-se de dois aspectos:

1. De que os arquivos e pastas que compõem as ferramentas CAPFITOGEN estejam alojadas em uma pasta e não se encontrem diretamente na raiz do diretório. Em outras palavras, o caminho para chegar a capfitogen.exe deveria ser X:\CAPFITOGEN\capfitogen.exe e não X:\capfitogen.exe, sendo X a letra da unidade de disco. Se necessário, reinstale as ferramentas assegurando-se cumprir a condição anteriormente detalhada.
2. De que no caso em que a pasta CAPFITOGEN não se encontre na raiz do diretório (X:\CAPFITOGEN), e sim num caminho dentro de outras pastas, algo não recomendável, este caminho não deve conter espaços entre palavras. Por exemplo, colocar a pasta no caminho X:\Mis Documentos\CAPFITOGEN irá gerar este erro pelo espaço que há entre as palavras “Mis” e “Documentos”.

Se apesar de assegurar-se de que as duas condições anteriores estejam cumpridas e o problema se repetir, ponha-se em contato com a assistência técnica do Programa.

n) Quando o processo que se executa depois de clicar no botão “analisar” leva muito tempo (mais de 15 ou 20 minutos) é possível que a página que mostra o formulário da ferramenta mude sua aparência e mostre um aviso de erro do java, como o que aparece na Fig. 12. Esta mudança e este aviso de erro não indicam que o processo tenha finalizado abruptamente nem que não se produziram

resultados. Se trata de um erro mais de visualização da interface que do processo em R. Por isso, o passo seguinte é revisar se aparecem os arquivos esperados como resultados na pasta designada para guardá-los.



Figura 12. Aviso de erro do java que aparece quando os processos levam tempo demais.

3. Ferramenta GEOQUAL

3.1. O que é a Avaliação da Qualidade da Georreferenciação nos dados de Passaporte?

É uma metodologia que determina o grau de certeza contido em alguns descritores de passaporte cuja função é definir inequivocamente o lugar onde o germoplasma foi coletado. Desta maneira, GEOQUAL faz uma avaliação da qualidade dos dados de descrição da localidade e das coordenadas indicadas como lugar de coleta.

Em termos gerais, o conceito de qualidade aplicado a dados recebeu diversas definições. No âmbito geográfico, a definição de qualidade como “aptidão de uso” ou potencialidade de uso tem sido amplamente aceita (Chrisman, 1983). Isto relaciona diretamente qualidade e possibilidade de uso dos dados. A incerteza associada a todo tipo de dados, é uma propriedade de quem obtém ou usa os dados mais que dos próprios dados. Por isso, qualidade e incerteza têm um grau de subjetividade variável, a qual pode reduzir-se até certo ponto mediante o uso de metodologias que realizem avaliações sobre bases o mais objetivas possíveis. Em todo caso, qualidade e incerteza se assumem como medidas de risco entendido e risco assumido (Chapman, 2005).

A necessidade de avaliar a qualidade da georreferenciação da informação disponível sobre a presença ou inclusive ausência de entidades biológicas, é um assunto palpável em diferentes áreas como a ecologia, análise espacial e padrões de distribuição de espécies. São numerosos os estudos que apontam que tal qualidade é um assunto crítico em metodologias como a modelação da distribuição de espécies. A certeza da ocorrência de uma espécie num local determinado é determinante para qualquer método que use dados de presença ou ausência como matéria-prima. Os seguintes estudos fazem referência a este aspecto: Foley e colaboradores, 2009; Hill e colaboradores, 2009; Otegui e colaboradores, 2013.

Ter uma estimativa do grau de incerteza na georreferenciação de locais de presença ou ausência de espécies se converte, assim, num aspecto chave antes de realizar qualquer análise que utilize o aspecto espacial para o estudo das distribuições. Muitas análises deste tipo desembocam na tomada de decisões em aspectos práticos de áreas como a conservação da biodiversidade. Portanto, introduzir informação de base confiável alimentando análises adequadas produzirá resultados confiáveis e decisões acertadas e oportunas.

3.2. Antecedentes da ferramenta GEOQUAL

A metodologia GEOQUAL acumula quatro anos de desenvolvimento, desde que surgiu a necessidade de um estimador de confiança (ou de risco) da georreferenciação do local de coleta, que normalmente ficam refletidos nos dados do passaporte. Em particular, a necessidade surgiu no fim do ano de 2009, no momento de preparar os dados do passaporte do Inventário Nacional de Recursos Fitogenéticos Espanhóis para serem caracterizados ecogeograficamente. Ter uma ideia da qualidade da georreferenciação em dados do passaporte foi então uma prioridade na hora de criar o Sistema de Informação Ecogeográfica dos Recursos Fitogenéticos Espanhóis (projeto SIERFE, <http://www.sierfe.es>).

SIERFE é um sistema que permite selecionar germoplasma a partir da caracterização ambiental do local de coleta através de um portal da internet. Ao desenvolver GEOQUAL e incorporá-lo no SIERFE, um estimador de qualidade permite aos usuários de SIERFE (solicitantes de germoplasma, como melhoradores, científicos ou agricultores) determinar sua exigência em termos de qualidade da georreferenciação na hora de selecionar germoplasma por variáveis ecogeográficas. Isto representa um avanço único no desenvolvimento de sistemas de informação e seleção de germoplasma. Mais de 45.000 entradas do Inventário Espanhol foram caracterizadas ecogeograficamente e cada uma recebeu um valor de qualidade em uma escala de 0 a 100.

GEOQUAL foi então desenvolvido especificamente para as características dos dados de passaporte do Inventário Nacional de Recursos Fitogenéticos Espanhóis e foi possível utilizando diversos programas, a maior parte de tipo comercial, como ArcGIS de ESRI.

Em 2011, no marco do projeto *PGR secure* do Sétimo Programa Marco da União Europeia (<http://www.pgrsecure.org>), foi necessária a depuração de quatro bases de dados que continham informação de ocorrência de variedades e espécies silvestres relacionadas com quatro gêneros de interesse agrícola na Europa (*Avena*, *Beta*, *Brassica* e *Medicago*). Mais de 33.000 registros receberam um valor GEOQUAL, o qual permitiu desconsiderar ou melhorar a qualidade de cerca de 4.000 registros.

A partir de então, vários pesquisadores europeus que trabalham em temas relacionados à agrobiodiversidade se interessaram por GEOQUAL, demandando o desenvolvimento de uma ferramenta de fácil manejo que permitisse aplicar GEOQUAL sobre diferentes formatos de dados de presença de espécies.

Já em 2012, quando o programa CAPFITOGEN é aprovado e se decidem as ferramentas a desenvolver, GEOQUAL se converteu em uma prioridade. Se tratava de abordar o desafio de criar uma ferramenta de avaliação da qualidade de dados da georreferenciação que fosse simples, com toda a informação necessária pré-carregada, que não requiera maiores conhecimentos em Sistemas de Informação Geográfica (SIG), para poder aplicá-la, que viesse em uma solução integrada (usando só um programa de SIG), que usasse como base o formato de

descritores de passaporte definidos pela FAO e Bioversity International em 2012 e que fosse transferível a técnicos de programas nacionais. A ferramenta GEOQUAL que se apresenta aqui, é portanto, a avaliação de uma ideia original transformada em tecnologia de fácil adoção, reunindo fatores de adaptabilidade apropriados às condições e necessidades de vários programas nacionais de conservação de recursos fitogenéticos.

3.3. Características de GEOQUAL

A ferramenta GEOQUAL se compõe de quatro parâmetros, três deles informam aproximações diferentes sobre a qualidade da georreferenciação (COORQUAL, SUITQUAL e LOCALQUAL) e um parâmetro final (TOTALQUAL) que resume os três primeiros. Os parâmetros-base se calculam em categorias que vão de zero a vinte, sendo zero, qualidade nula e 20, máxima qualidade. Em algumas ocasiões, dependendo dos dados de passaporte disponíveis, se pode prever o cálculo de LOCALQUAL, como se explica mais adiante. Adicionalmente se gera um parâmetro que transforma os valores iniciais de TOTALQUAL (0 a 40 ou 0 a 60) a uma categoria de avaliação de 0 a 100, para facilidade de uso e interpretação do valor de avaliação (TOTALQUAL100).

É importante ter em conta que GEOQUAL opera sobre o formato de descritores de passaporte FAO-Bioversity publicado no ano 2012 com o acréscimo de quatro descritores de localidade (ADM1, ADM2, ADM3 e ADM4) que têm correspondência com diferentes figuras administrativas segundo cada país (ver Anexo 9.1). No entanto, se os dados estiverem no formato FAO-IPGRI 2001, GEOQUAL poderia também operar prévio traslado da informação do formato 2001 ao 2012 sem ter que acrescentar informação para os novos campos que inclui a versão 2012, mas sim considerando a inclusão dos quatro descritores ADM.

A ferramenta GEOQUAL inclui um modelo de tabela de descritores de passaporte baseado nos descritores multi-cultivo de FAO-Bioversity com o acréscimo dos quatro descritores ADM em formato Excel (pasta “DescriptoresPasaporteModelo”, arquivo “Tabla pasaporte modelo FAO_Bioversity 2012 modificada.xls”), onde em cor verde se assinala os descritores indispensáveis para GEOQUAL e, em cor amarela, aqueles não indispensáveis, mas sim, importantes. Os campos não assinalados não são levados em conta por GEOQUAL, mas deve-se manter sua posição dentro da tabela (como no caso dos assinalados) para que GEOQUAL encontre as variáveis que analisa no local que espera encontrá-las. Como regra geral, ao preencher esta tabela, quando não conheça a informação que se requer, se deve escrever no campo as letras NA, que normalmente indicam “não aplica” mas que no caso de GEOQUAL também indica que não há informação disponível.

3.3.1 Descrição dos parâmetros base de GEOQUAL

3.3.1.1 Parâmetro COORQUAL

É um parâmetro que determina a qualidade intrínseca das coordenadas contidas nos dados do passaporte. Se determina através de quatro sub-parâmetros inicialmente:

a) ERRORES: Se as coordenadas em formato sexagesimal ou decimal contêm valores fora do marco possível dentro do sistema de coordenadas latlong WGS84. Usa os descritores LATITUDE, LONGITUDE, DECLATITUDE e DECLONGITUDE.

b) PRECIS: Aplica para coordenadas em formato sexagesimal que cumpram a codificação do listado de descritores do passaporte FAO-Bioversity 2012. Este sub-parâmetro determina se as coordenadas foram obtidas com uma precisão de segundos, minutos ou graus. Usa os descritores

c) GEORBLE: É uma valorização da possibilidade de obter coordenadas do local de coleta a partir dos dados de descrição de localidade disponíveis.

d) INTERTEMP: Toma valores do descritor COLLDATE e os interpreta na medida da possibilidade do uso de métodos de georreferenciação. Por exemplo, para coletas ocorridas posteriormente ao ano 2000, se assume uma alta probabilidade de uso de GPS, o qual incrementaria a qualidade das coordenadas.

e) GEOREFMETH: Valoriza o sistema pelo qual se tem atribuído coordenadas ao local de coleta. GEOREFMETH corresponde a um campo da tabela do passaporte FAO/Bioversity 2012. Somente se considerará este sub-parâmetro quando se dispõe de valores para todas as entradas deste campo.

Cada sub-parâmetro se avalia em uma categoria de zero a três, onde zero corresponde a qualidade mínima e três a qualidade máxima. As combinações dos valores de cada sub-parâmetro geram o parâmetro COORQUAL numa categoria de zero a vinte.

3.3.1.2 Parâmetro SUITQUAL

Este parâmetro atribui um valor de qualidade às coordenadas de acordo com o apropriado do local de coleta para o crescimento de plantas, e diferenciando da natureza da entrada (silvestre ou cultivada de acordo com o descritor SAMPSTAT). A informação sobre as características do local da coleta provém de um mapa de uso do solo (Global Land Cover 2000 ou GLC2000). Se trata do mapa de cobertura global de livre acesso sobre uso do solo mais antigo, que conta com uma resolução apropriada (1 km). As classes originais deste mapa se transformam em função do apropriado que resulta cada classe à presença de plantas cultivadas ou silvestres, em uma escala de 0 a 20.

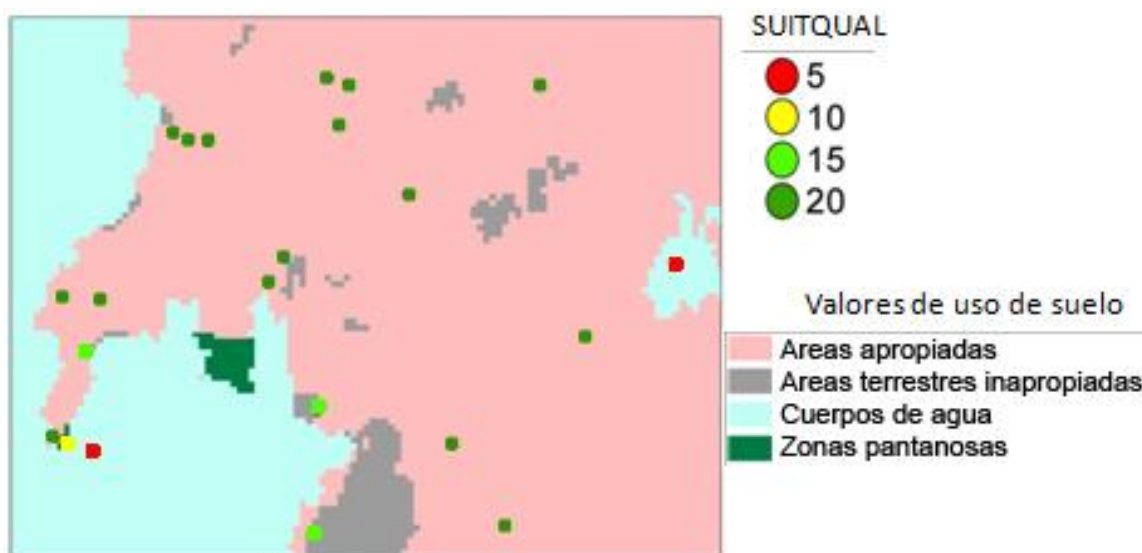


Figura 13. Exemplo da obtenção de valores SUITQUAL de acordo a interpretação dos valores de uso do solo.

3.3.1.3 Parâmetro LOCALQUAL

LOCALQUAL é o resultado da comparação da descrição da localidade onde se coletou o germoplasma procedente dos campos ORIGCTY, ADM1, ADM2, ADM3, ADM4 e COLLSITE, com os campos ISO, NAME1, NAME2, NAME3 e NAME4 da base de dados “Global Administrative Areas” (GADM) v2.0, extraídos por meio das coordenadas provistas em DECLATITUDE e DECLONGITUDE (ou através da transformação ao formato decimal de LATITUDE e LONGITUDE). A diferença do processo que realiza “Check Coordinates” (comprovar coordenadas) incluído em DIVA-GIS onde a comparação é absoluta (os termos devem comparar característica por característica, buscando a coincidência), GEOQUAL utiliza a distância generalizada de Levenshtein através da função “agrep” do pacote base de R, que toma em conta o número de inserções, eliminações ou mudança de característica entre as duas cadeias que se comparam. Assim, permitindo um certo número deste tipo de mudanças, a função “agrep” é capaz de identificar concordâncias ainda que ocorram erros de tipografia ou pelo efeito daqueles caracteres alfabéticos próprios de certos idiomas, que em ocasiones não se codificam corretamente (caso da ñ e os acentos no idioma castelhano).

Para mais segurança, LOCALQUAL compara também com os campos VARNAME1, VARNAME2, VARNAME3 e VARNAME4 incluídos em GADM, que são variantes do nome oficial da unidade administrativa e que podem ser usados pelos curadores na hora de registrar o germoplasma em suas bases de dados de passaporte.

Finalmente, LOCALQUAL considera a série de comparações positivas entre os diferentes emparelhamentos (ORIGCTY com ISO, ADM1 com NAME1, etc.) para calcular um valor numa categoria de zero a vinte.

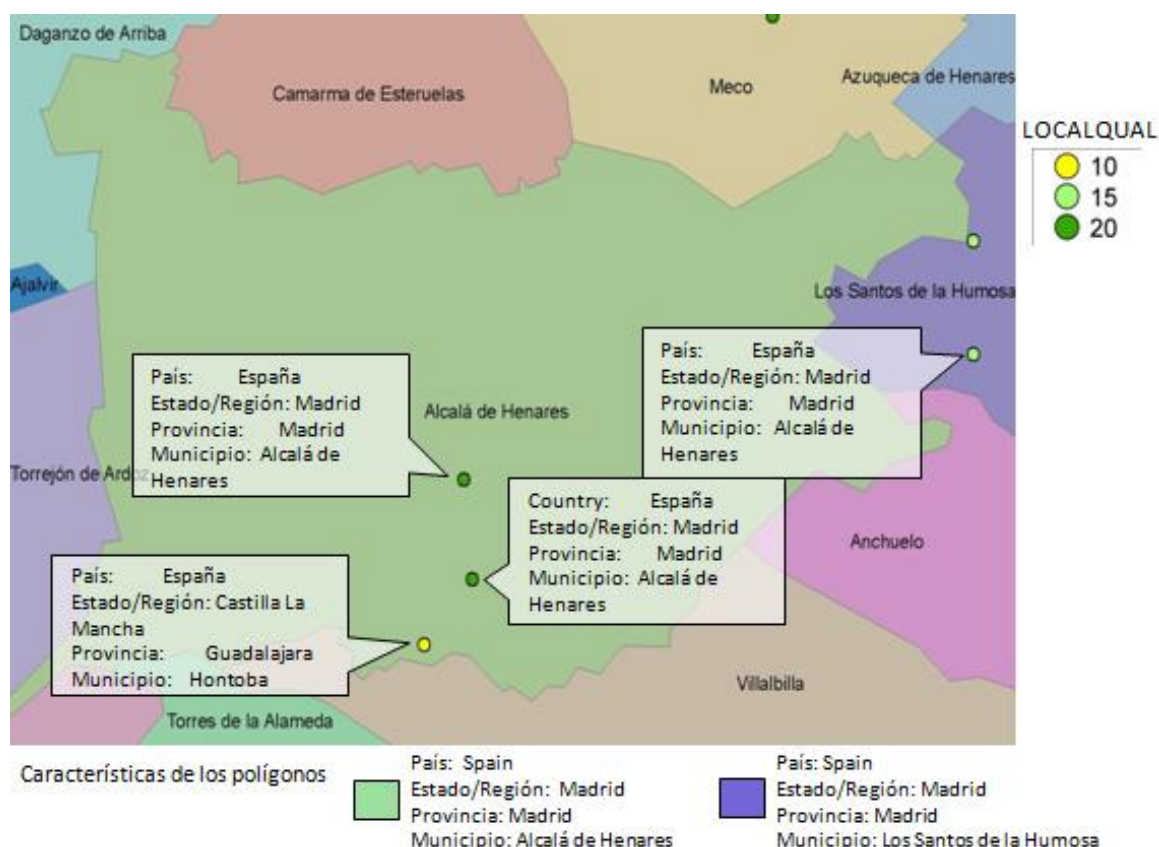


Figura 14. Exemplo da obtenção de valores LOCALQUAL de acordo com a comparação de níveis administrativos dos dados informados pelo usuário e os extraídos pelas coordenadas a partir de GADM.

3.3.2 Descrição dos parâmetros TOTALQUAL e TOTALQUAL100

O parâmetro final do resumo TOTALQUAL é simplesmente a soma dos valores de COORDQUAL, SUITQUAL e LOCALQUAL. Pelas categorias possíveis de valores destes três parâmetros, TOTALQUAL pode tomar valores entre 0 e 60. No entanto, para facilidade de interpretação e análise dos resultados de GEOQUAL, se calcula também o parâmetro TOTALQUAL100 que é uma transformação de TOTALQUAL a uma categoria de valores de 0 a 100, onde 0 é qualidade nula (incluindo a falta de coordenadas) e 100 representa uma qualidade teórica ótima.

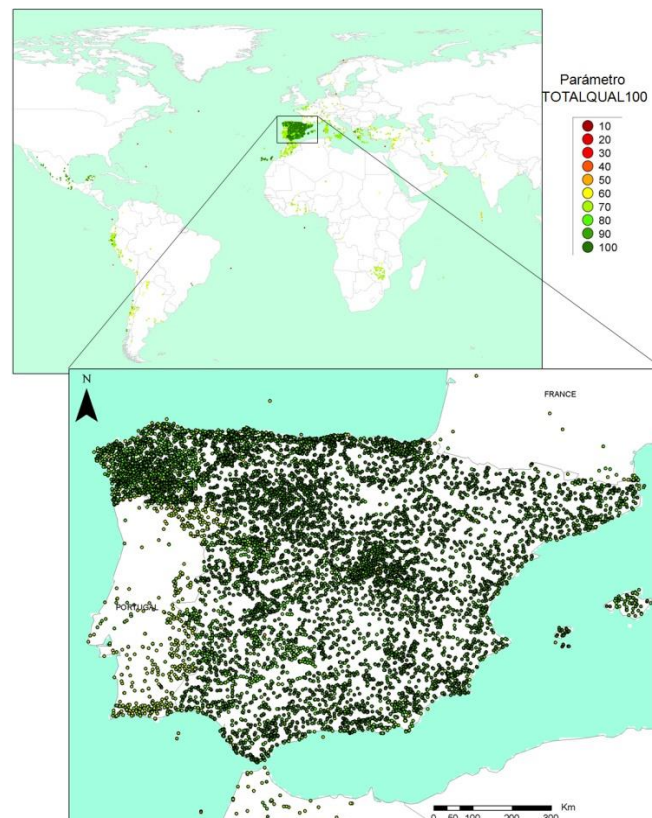


Figura 15. Resultados da aplicação de GEOQUAL no Inventário Nacional Espanhol de Recursos Fitogenéticos. Se mostram os valores alcançados de TOTALQUAL100.

3.3.3 Determinação de limites de qualidade

GEOQUAL tem sido pensada desde sua implementação inicial, como uma metodologia altamente objetiva, na qual o usuário intervém minimamente na obtenção do valor definitivo. No entanto, toda determinação de qualidade implica componentes subjetivos e GEOQUAL não é exceção.

Por exemplo, existe certa subjetividade quando se aplica a certas categorias de uso do solo os valores de aptidão para o crescimento de plantas. Também a definição de a partir de que valores se pode considerar alta ou baixa qualidade é uma questão subjetiva que tem mais a ver com o observador do que com a técnica.

O limite sobre o qual se considera que uma entrada está corretamente georreferenciada usando valores GEOQUAL deve ser definida pelo usuário dos dados com base em suas expectativas e necessidades. Pode-se pensar o uso de diferentes limites dependendo do uso que se dará aos dados, do tipo de estudo a que serão submetidos, e da precisão e exatidão da informação do resto das fontes num determinado estudo. É recomendável ver a distribuição de valores TOTALQUAL100 do conjunto de entradas e assim saber de antemão que ao aplicar um limite muito exigente (próximo a 100) ou pouco exigente (abaixo de 50) se selecionariam muitas ou poucas entradas.

3.4. Utilização da ferramenta GEOQUAL

Uma vez instalada as ferramentas CAPFITOGEN e selecionada a ferramenta GEOQUAL, será necessário definir uma série de parâmetros para que a programação R funcione corretamente.

Depois de definir todos os parâmetros e caminhos que requer GEOQUAL, ao clicar no botão “Analisar”, se iniciará o processo de análise da ferramenta. Então, depois de um tempo que pode variar pela introdução de alguns parâmetros de resolução, tipo de análise, tamanho de dados processados ou configuração de hardware do computador, GEOQUAL produzirá resultados os quais guardará onde tenha sido indicado (parâmetro 3.4.1.5).

3.4.1 Parâmetros iniciais definidos pelo usuário

3.4.1.1 Parâmetro: ruta

Explicação: Caminho é onde se encontram as ferramentas (documentos) CAPFITOGEN. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo: F:/CAPFITOGEN, C:/CAPFITOGEN, D:/MinhasFerramentas (MeusDocumentos)/CAPFITOGEN, etc.

3.4.1.2 Parâmetro: pasaporte

Explicação: Escrever o nome do arquivo que contém a tabela de passaporte em formato de texto, lembrando de incluir a extensão (.txt). Por exemplo, se o arquivo se chama "tabela" deve-se escrever: "tabela.txt". Recorde que este arquivo deverá ser guardado previamente na pasta "Passaporte", que faz parte do conjunto de pastas que compõe o diretório CAPFITOGEN.

3.4.1.3 Parâmetro: precision

Explicação: Escolher usar mapas de alta ou baixa precisão para determinar se as coordenadas de um lugar de coleta caem no mar e em qual distancia. Alta resolução poderá desacelerar um pouco o processo quando a base de dados for muito grande (superior a 15.000 entradas com coordenadas).

3.4.1.4 Parâmetro: local

Explicação: Indicar se deseja usar o parâmetro LOCALQUAL para avaliação da qualidade do georeferenciamento. LOCALQUAL é um parâmetro de comparação entre o local descrito e o extraído pelo SIG. Se seus dados não contêm nenhuma descrição de local ou a descrição está contida por completo no campo COLLSITE, NÃO é conveniente usar esta opção.

3.4.1.5 Parâmetro: resultados

Explicação: Introduzir o caminho da pasta onde deseja guardar os resultados das análises. Nota: usar / em vez de \ na indicação do caminho da pasta. Por

exemplo C:/Resultados, D:/MinhasFerramentas(Meus Documentos)
/Resultados, etc.

3.5. Resultados de GEOQUAL

No caminho e pasta determinada em “resultados” (parâmetro 3.4.1.5) se encontrará três tabelas e um mapa de pontos do tipo vetorial (shapefile).

3.5.1 Tabelas

As tabelas produzidas por GEOQUAL estão em formato de texto delimitado por tabulações e podem ser abertas em programas como Excel, OpenOffice ou R.

3.5.1.1 “PasaporteOriginalEvaluadoGEOQUAL.txt”: É a tabela de passaporte no formato sugerido que originalmente se usou para a análise, com acréscimo de cinco colunas com os valores obtidos para os parâmetros: SUITQUAL, LOCALQUAL, COORQUAL, TOTALQUAL e TOTALQUAL100.

3.5.1.2 “tabla_de_analisisGEOQUAL.txt”: Esta tabela contém também todas as colunas da tabela de passaportes originalmente introduzidas para sua análise, ainda que neste caso só se incluem as entradas com coordenadas. No entanto, o mais importante desta tabela é que nela foram acrescentadas todas as colunas correspondentes a extrações, interpretações ou subparâmetros que foram necessários para poder calcular os valores dos parâmetros de GEOQUAL. A lista de variáveis adicionais incluídas nesta tabela e sua explicação se encontra no Anexo 12.4.

3.5.2 Mapas

3.5.2.1 Mapa de pontos em formato vetorial tipo “shapefile”. Este mapa vem acompanhado de uma tabela que inclui os valores dos parâmetros de avaliação GEOQUAL, de tal maneira que em DIVA-GIS se podem mostrar os pontos com diferentes graus de cor de acordo a sua pontuação (qualidade). Um “shapefile” se compõe de até 6 arquivos de igual nome mas diferente extensão. No caso de GEOQUAL, o shapefile só está composto por três extensões (.shp, .shx, e .dbf) e tem o nome de ShapefilePuntosGEOQUAL.

3.5.2.2 Mapa de pontos em formato Google Earth. Este mapa corresponde ao arquivo mapa_puntos_google.kml e se tiver instalado o programa Google Earth. Clicando duas vezes em seu nome no Windows Explorer, o mapa de pontos (em forma de alfinetes ou tachinhas) se abrirá neste mesmo programa, localizando os locais de coleta sobre as imagens de satélite. Ao clicar sobre os alfinetes se abrirá uma pequena janela mostrando o valor TOTALQUAL100 de cada entrada.

3.6. Referências

Chapman, A.D. 2005. Principles of data quality, versión 1.0. Report of the Global Biodiversity Information Facility, Copenhagen.

Chrisman, N.R. 1983. The role of quality information in the long-term functioning of a GIS. Proceedings of AUTOCART06, 2: 303-321. Falls Church, VA: ASPRS.

FAO, IPGRI. 2001. Lista de descriptores de pasaporte para cultivos múltiples desarrollada por la FAO y el IPGRI .

FAO, BIOVERSITY. 2012. FAO/Bioversity multi-crop Passport descriptors V.2. Disponible en http://www.biodiversityinternational.org/index.php?id=19&user_biodiversitypublications_pi1%5BshowUid%5D=6901

Foley, D.H., Wilkerson, R.C., Rueda, L.M. 2009. Importance of the "what," "when," and "where" of mosquito collection events. J Med Entomol. 2009 Jul;46(4):717-22.

Hill, A.W., Guralnick, R., Flemons, P., Beaman, R., Wieczorek, J., Ranipeta, A., Chavan, V., Remsen, D. 2009. Location, location, location: utilizing pipelines and services to more effectively georeference the world's biodiversity data. BMC Bioinformatics. 2009 Nov 10;10 Suppl 14:S3. doi: 10.1186/1471-2105-10-S14-S3.

Otegui, J., Ariño, A.H., Encinas, M.A., Pando, F. 2013. Assessing the primary data hosted by the Spanish node of the Global Biodiversity Information Facility (GBIF). PLoS One. 2013;8(1):e55144. doi: 10.1371/journal.pone.0055144.

Soberón, J., Peterson, T. 2004. Biodiversity informatics: managing and applying primary biodiversity data. Phil. Trans. R. Soc. Lond. B. 359, 689-698.

4. Ferramenta ELC mapas

4.1. O que é um mapa de caracterização ecogeográfica do terreno?

Trata-se de um tipo de mapa onde podemos visualizar diferentes cenários ambientais que podem se corresponder com os diferentes processos adaptativos de uma espécie vegetal ao longo de um determinado território. Entre seus múltiplos usos, os mapas de caracterização ecogeográfica do terreno (mapas ELC) são úteis para a conservação e uso razoável da agrobiodiversidade.

A ideia de expressar adaptação através de mapas não é nova. Desde meados do século passado se vem desenvolvendo mapas de biomas, ecossistemas, regiões ecológicas, etc. Estes mapas representam unidades ambientais geralmente como regiões grandes e homogêneas. Em princípio os “climas” ou “ambientes” (os termos eram usados indistintamente) que representam estes mapas eram usados em estudos de diversos tipos de organismos (plantas, animais, microorganismos). Alguns mapas afinaram mais e representavam, por exemplo, os climas afins às formações vegetais descritas por Leslie Holdridge em 1947, ainda que logo se generalizaram como “sistemas de classificação de zonas de vida”.

Estes mapas foram de muita utilidade para biólogos e naturalistas que tentavam compreender a distribuição de organismos vivos em relação à temperatura e a humidade. Ainda hoje em dia, o sistema de Holdridge tem aplicação, por exemplo, em estudos de mudança climática. Contudo, a mistura às vezes indistinguível de características bióticas (vegetação) e abióticas (temperatura, precipitação) nestes mapas, a simplificação do componente abiótico em só dois fatores e a forma de delinear as regiões (grandes, homogêneas e contínuas) representavam um impedimento sério para seu uso em estudos de adaptação em espécies em particular.

A utilização da informação sobre adaptação para desenhar uma coleção, ou ter critérios na hora de conservar e utilizar de maneira eficiente os recursos genéticos vegetais tão pouco é algo novo, ainda que são escassos os trabalhos publicados sobre a matéria onde explicitamente se faça referência à adaptação. Existe, como referência, um primeiro mapa ecogeográfico cujo propósito era criar coleções núcleo/nucleares (core collections) em 1997 (Tohme et al., 1995), onde ao final do processo de seleção de entradas, além da dimensão ecogeográfica, se consideraram outros critérios adicionais de diferente natureza.

Desde então, tem-se produzido vários desenvolvimentos: os programas SIG tornaram-se gradualmente mais flexíveis e “amigáveis” e, inclusive, alguns pacotes estatísticos, atualmente, incluem utilidades e ferramentas SIG; a informação ecogeográfica disponível (em forma de capas SIG) é de melhor qualidade e mais acessível; os equipamentos de informática de alta capacidade de análise são vendidos a preços reduzidos e o acesso à internet tem-se

incrementado em países em desenvolvimento. Estes avanços têm repercutido no desenvolvimento de mapas que representam diferentes cenários de adaptação para plantas cultivadas e silvestres aparentadas. Foi assim como se obteve um primeiro mapa de caracterização ecogeográfica do território para Espanha em 2005 (Parra Quijano et al., 2008) de tipo generalista (que pudesse ser aplicado a várias espécies silvestres aparentadas com cultivadas) ainda que só foi utilizado para algumas espécies do gênero *Lupinus*. Este mapa, obtido por técnicas de análise multivariada e determinação do número de grupos por critérios bayesianos, representou as diferentes unidades ambientais de forma reticulada e como regiões homogêneas, mas de dimensões reduzidas e descontínuas. Estas características físicas já contrastavam com os mapas bioclimáticos tradicionais. Outra diferença foi a inclusão de variáveis de tipo geofísico e edáfico (além das bioclimáticas), com a ideia de representar os aspectos abióticos que podem afetar o desenvolvimento de uma planta desde um ponto de vista agrônomo.

A meados de 2008 se desenvolveu um novo mapa ecogeográfico para Espanha Peninsular e Ilhas Baleares usando outras fontes de informação ecogeográfica. A metodologia de criação deste novo mapa foi similar ao anterior. Neste ponto, o que mais interessava aos pesquisadores era verificar se o mapa realmente refletia cenários adaptativos, ou seja, realizar uma validação. Para isso, se avaliou o desempenho do novo mapa para oito espécies (quatro leguminosas e quatro gramíneas), duas delas silvestres aparentadas com cultivadas e as outras seis compostas por variedades locais, usando sua distribuição, sendo a variável “peso da semente” como variável fenotípica indicadora com valor adaptativo. Os resultados se compararam com dois mapas de referência, um com uma estrutura física similar ao mapa ecogeográfico (alta reticulação, descontinuidade, unidades de tamanho reduzido), mas construído sem ter em conta aspectos relacionados com a adaptação abiótica de plantas (mapa CORINE land cover, um mapa de uso do solo, ver <http://www.eea.europa.eu/data-and-maps/explore-interactive-maps/corine-landcover-2006>) e outro com uma estrutura física diferente (mais parecida a dos mapas tradicionais) mas construído com um fim similar (mapa DMEER ou mapa digital de regiões ecológicas europeias, ver <http://www.eea.europa.eu/data-and-maps/figures/dmeer-digital-map-of-european-ecological-regions>).

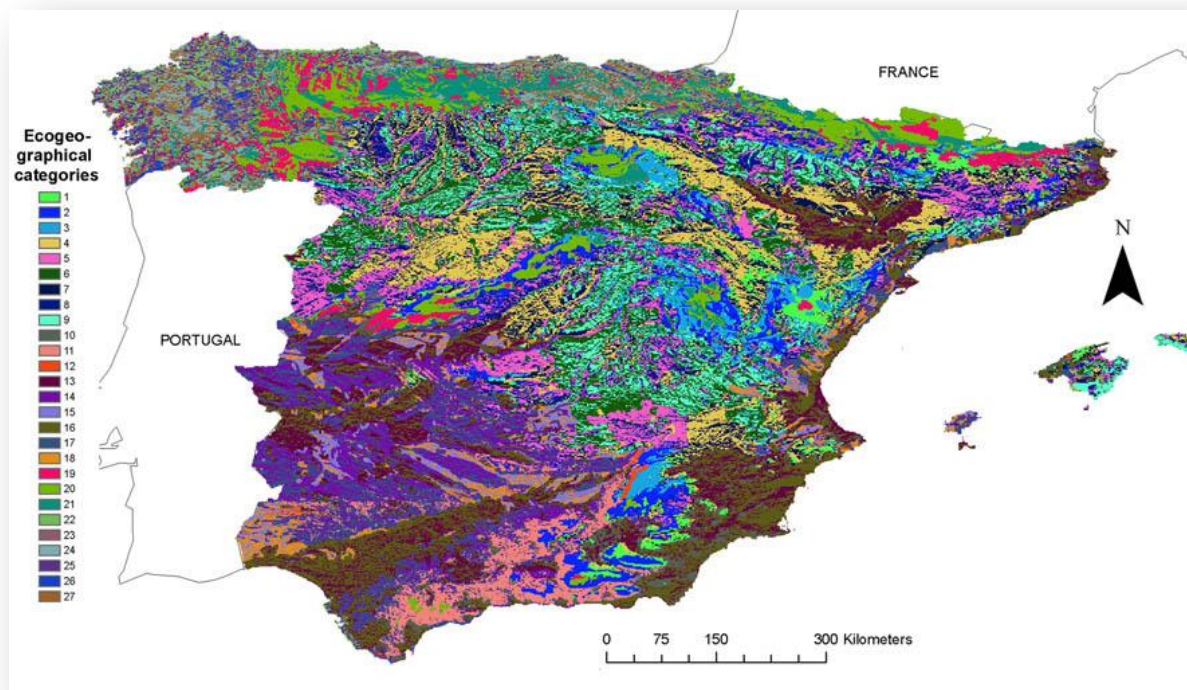


Figura 16. . Mapa de caracterização ecogeográfica do território da Espanha Peninsular e Baleares para *Lupinus*.

Os resultados do estudo de validação foram variados. O mapa de caracterização ecogeográfica do território funcionou em geral melhor para espécies leguminosas que para gramíneas, ainda que a exceção foi *Zea mays* para a qual o desempenho do mapa foi bastante aceitável. Como era de se esperar, o mapa refletiu cenários adaptativos para as duas espécies silvestres, mas também produziu resultados satisfatórios nos casos de espécies compostas só por variedades locais, como o caso de *Phaseolus vulgaris*. Como conclusão, os mapas de caracterização ecogeográfica do território são capazes de refletir cenários adaptativos e, portanto, podem ser utilizados em muitas atividades relacionadas com a coleta, conservação e utilização eficiente dos recursos genéticos vegetais. Contudo, é recomendável criar mapas específicos para cada espécie ou grupo de espécies relacionados filogeneticamente. Realizar mapas de caracterização ecogeográfica do terreno de tipo generalista pode resultar arriscado na hora de tirar conclusões para um grupo numeroso de espécies, particularmente se não se realiza uma validação do mapa. Também é importante fazer uma seleção apropriada das variáveis ecogeográficas envolvidas na criação do mapa, mas sempre representando os três aspectos abióticos envolvidos no desenvolvimento das plantas: bioclimático, geofísico e edáfico.

4.2. Antecedentes da ferramenta ELC mapas

A ferramenta ELC mapas que aborda esta instrução representa o desenvolvimento do conceito *Ecogeographical Land Characterization Maps* exposto na publicação de Parra Quijano e colaboradores (2012 A)

Este tipo de mapas tem tido usos diversos em coleta, conservação e uso de recursos fitogenéticos como nos trabalhos de Parra Quijano e colaboradores (2011 A, 2011 B e 2012 B) ou Thormann (2012).

O interesse que esta metodologia havia despertado em várias equipes e projetos de pesquisa em coleta, conservação e uso de recursos genéticos vegetais contrastava de uma observação recorrente entre possíveis usuários: a metodologia descrita em tal publicação é um tanto complexa dado que mistura técnicas de sistemas de informação geográfica (SIG) e análise multivariada. Além disso, o desenvolvimento original implicava o uso de um programa de análise estatística pago. Estes aspectos limitavam de maneira importante o desenvolvimento de mapas ecogeográficos de caracterização do terreno por parte de pesquisadores e técnicos.

4.3. Características de ELC mapas

A ferramenta ELC mapas provê uma nova opção para desenvolver mapas de caracterização ecogeográfica do terreno mediante o uso de R, evitando as complicações anteriormente descritas. Este software, gratuito e com ampla capacidade de cálculo estatístico e potência gráfica, permite integrar SIG e análise multivariada. Assim, a nova ferramenta ELC mapas é capaz de produzir mapas de caracterização ecogeográfica do terreno sem necessidade de alternar entre diferentes programas, nem descarregar e manipular informação ecogeográfica. Contudo, é importante esclarecer que os produtos da ferramenta ELC mapas são mapas e tabelas que podem ser visualizadas em programas como DIVA-GIS, Google Earth ou Microsoft Excel. Os mapas de caracterização ecogeográfica do terreno da ferramenta ELC mapas são assim mesmo um componente de outras ferramentas como Representa.

ELC mapas usa dois procedimentos para determinar o número de grupos a usar dentro do método de agrupamento. Estes procedimentos são:

- a) Um sistema simples que usa como algoritmo de agrupamento K –means com determinação do ponto de corte baseado na diminuição da soma de quadrados intragrupo (Ketchen e Shook, 1996). O número de grupos ótimo é o alcançado quando a diminuição da somas de quadrados intragrupos entre uma solução de n e uma de $n+1$ grupos é menor que 50%. Este método, também conhecido como “elbow” ou cotovelo, é o de mais rápido cálculo e é capaz de atuar sobre grandes quantidades de dados sem grandes demoras, e por isso se aconselha para países de grande tamanho.

b) Método de agrupamento de partição em torno dos medoides (pam). Utiliza-se o método de silhouette de interpretação e validação do número de grupos. Este sistema (em princípio gráfico, posteriormente adaptado em R pelo pacote fpc) permite determinar o quão apropriadamente tem sido os dados agrupados (Kaufman e Rousseeuw, 1987; Rousseeuw, 1987). É um sistema que consome mais recursos computacionais, e por isso, se aplicado a grandes conjuntos de dados, consumirá mais tempo.

Os métodos de determinação do número de grupos não são de todo objetivos, pois, minimamente, o usuário deve determinar um número máximo de grupos que deseja e, para o método do cotovelo, o percentual de diminuição é subjetivo, ainda que baseado na observação dos gráficos de variação intragrupos por número de grupos.

Dado que a informação ecogeográfica a níveis de resolução de 1 km ou inclusive de 5 km é bastante volumosa para um subcontinente completo como América Latina, a ferramenta ELC mapas está inicialmente pensada para ser utilizada a nível de país, se bem que a distribuição das espécies objetivo ou a distribuição das coletas de germoplasma possa ir mais além dos limites de um país determinado. Se espera produzir, em um curto prazo, uma versão de ELC mapas com dados a nível continental ou subcontinental com menos resolução (10 km).

4.4. A seleção das variáveis ecogeográficas

Este é um aspecto muito importante a ter em conta antes de usar a ferramenta ELC mapas. A mudança, acréscimo ou eliminação de uma só variável de um só componente (bioclimático, geofísico ou edáfico) mudará sensivelmente a configuração final do mapa e sua correlação com os cenários adaptativos da espécie.

A técnica de mapas de caracterização ecogeográfica do terreno originalmente não contemplava fazer maior discriminação de variáveis, uma vez que se pensava em mapas de corte generalistas. Contudo, à medida que se reconheceu que sua capacidade de discriminar corretamente cenários adaptativos se incrementava ao focar-se em uma espécie em particular ou um grupo de espécies muito afins (em termos genéticos), surgiu a ideia de selecionar as variáveis ecogeográficas de cada componente que mais influenciam na adaptação abiótica da espécie ou espécies, e que, portanto, determinam sua distribuição.

O processo de seleção é, então, um aspecto chave na hora de obter mapas mais precisos em termos adaptativos. O listado de variáveis potencialmente selecionáveis se pode criar a partir de:

a) Pesquisas bibliográficas: Não é difícil encontrar na literatura de tipo técnica e/ou científica, como artigos, livros ou documentos, algumas referências sobre os fatores ambientais que influenciam, determinam ou

limitam a distribuição de uma espécie. Frequentemente deve-se buscar a correspondência entre as referências sobre esses fatores com as variáveis disponíveis em forma de capas SIG para realizar o mapa.

b) Conhecimento especializado: A consulta de pessoas especialistas na espécie ou grupo de espécies, sobre quais variáveis ecogeográficas podem ser consideradas fatores chave na sua adaptação e distribuição costuma oferecer uma informação muito valiosa na hora de selecionar variáveis. Esta consulta introduz subjetividade no processo, algo que não se deve temer em absoluto. No processo de criação de mapas de caracterização ecogeográfica do terreno, o conhecimento especialista nestas fases prévias pode ser a diferença entre um mapa corretamente validado ou um mapa com pouco sentido em termos de adaptação da espécie objetivo. O aporte do conhecimento especializado pode ser mais decisivo na medida que se consulte o maior número de expertos possível e logo se determinem consensos. Um bom exemplo de mapa de caracterização ecogeográfica do terreno em que foi utilizado conhecimento especializado, é o trabalho de Parra Quijano e colaboradores (2012 C).

Neste estudo, o mapa foi utilizado para a determinação da localização idônea de reservas genéticas para várias espécies do gênero *Beta* na Europa.

Posterior à elaboração deste listado potencial de variáveis, se deve determinar por cada componente (bioclimático, geofísico e edáfico) quais variáveis podem estar contribuindo com informação redundante. Para isso é importante realizar uma análise de correlações bivariadas ou uma análise de colinealidade. No caso de encontrar-se valores altos de correlação entre duas variáveis, que façam parte de um mesmo componente, se deveria descartar uma delas. Além disso, uma análise de componentes principais (no caso de que todas as variáveis sejam quantitativas) pode ajudar a determinar as relações das variáveis e facilitar a seleção final. Não é recomendável utilizar mais de cinco variáveis por componente, pois a configuração das zonas (células adjacentes com o mesmo valor) do mapa resultante, pode chegar a ser difícil de compreender. Assim mesmo, utilizar latitude e longitude (parâmetros 4.5.1.7 e 4.5.1.8) oferece mapas com zonas maiores e mapas menos reticulados. O efeito contrário se obtém ao utilizar variáveis como “orientação” do componente geofísico.

Uma vez determinada a lista final de variáveis, estas se selecionam nos parâmetros bioclimv, geophysv e edaphv (parâmetros 4.5.1.5, 4.5.1.6 e 4.1.5.9). A lista de variáveis completa que inclui a ferramenta ELC mapas v1.2 se encontra na parte final deste documento (Anexos 12.1, 12.2 e 12.3).

4.5. Utilização da ferramenta ELC mapas

Uma vez instaladas as ferramentas CAPFITOGEN e selecionada a ferramenta ELC mapas, deve-se especificar uma série de parâmetros por parte do usuário.

4.5.1 Parâmetros iniciais definidos pelo usuário

Depois de definir todos os parâmetros e caminhos que requer ELC mapas, ao clicar em “Analisa”, a ferramenta iniciará a análise. Depois de um tempo que pode variar pela introdução de alguns parâmetros de resolução, tipo de análise, tamanho de dados processados ou configuração de hardware do computador, ELC mapas produzirá resultados que guardará onde foi indicado (parâmetro 4.5.1.12).

4.5.1.1 Parâmetro: ruta

Explicação: Caminho é onde se encontram as ferramentas (documentos) CAPFITOGEN. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo: F:/CAPFITOGEN, C:/CAPFITOGEN, D:/MinhasFerramentas (MeusDocumentos)/CAPFITOGEN, etc.

4.5.1.2 Parâmetro: pais

Explicação: Selecionar o país para o qual se deseja construir o mapa ELC. Aparecerão vários países disponíveis somente se a ferramenta incluir informação para todos eles, se não, só aparecerá um determinado país escolhido.

4.5.1.3 Parâmetro: resol1

Explicação: Selecionar o nível de resolução que se deseja usar para gerar o mapa. Deve-se considerar que 1x1 km oferece melhor resolução, mas exige maior capacidade computacional e levará mais tempo do que 5x5 km, especialmente em países de grande extensão territorial. Ver o Anexo 12.5 sobre disponibilidade de resoluções em relação ao país ou região selecionada.

4.5.1.4 Parâmetro: bioclimv

Explicação: Selecionar as variáveis bioclimáticas (temperaturas, precipitações e índices) que se deseja incluir para gerar o mapa ELC. Podem-se selecionar múltiplas variáveis mantendo pressionada a tecla Ctrl (control) e adicionando mais variáveis clicando sobre elas com o botão esquerdo do mouse.

4.5.1.5 Parâmetro: geophysv

Explicação: Selecionar as variáveis geofísicas (relativas ao relevo e radiação solar) que se deseja incluir para gerar o mapa ELC. Podem-se selecionar múltiplas variáveis.

4.5.1.6 Parâmetro: latitud

Explicação: Incluir latitude em seu mapa? Nota: ao se incluir latitude e longitude e excluir orientação, se criarão unidades ecogeográficas mais contíguas (próximas) e mapas menos reticulados.

4.5.1.7 Parâmetro: longitud

Explicação: Incluir longitude em seu mapa? Nota: ao se incluir latitude e longitude e excluir orientação, se criarão unidades ecogeográficas mais contíguas (próximas) e mapas menos reticulados.

4.5.1.8 Parâmetro: edaphv

Explicação: Selecionar as variáveis edáficas (texturas de solo, profundidade, pH, etc.) que se deseja incluir para gerar o mapa ELC. Podem-se selecionar múltiplas variáveis.

4.5.1.9 Parâmetro: optim

Explicação: Indicar se requer processo de otimização. A otimização somente é recomendável para países de grande extensão territorial (Argentina, Brasil, México) usando alta resolução (1x1 km). Do contrario, pode-se desacelerar o processo antes que acelerá-lo. Além disso, se for incluído latitude ou longitude como variáveis para criar o mapa ELC, não se deve usar “optim”.

4.5.1.10 Parâmetro: maxg

Explicação: Indicar o número máximo de agrupamentos (grupos) por componente (bioclimático, geofísico e edáfico) que se deseja permitir (quanto maior o número, maior a quantidade de categorias por mapas). É recomendável que se use menos de 5 (cinco).

4.5.1.11 Parâmetro: metodo

Explicação: Selecionar um dos métodos oferecidos para gerar agrupamentos (grupos) com corte objetivo. “elbow” ou codo (cotovelo) é o método mais simples e rápido; “medoides” é o mais sofisticado e requer mais recursos.

4.5.1.12 Parâmetro: resultados

Explicação: Introduzir o caminho da pasta onde deseja guardar os resultados das análises. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo C:/Resultados, D:/MinhasFerramentas(Meus Documentos)/Resultados, etc.

4.6. Resultados de ELC mapas

No caminho e pasta determinada em “resultados” (parâmetro 4.5.1.12) se encontrará cinco mapas e três tabelas.

4.6.1 Os mapas

Correspondem ao mapa ELC final do país ou a região determinada no parâmetro 4.5.1.3 (mapa_elc_pais.grd, e mapa_elc_pais.gri, junto com a imagem mapa_elc_pais.png) e os mapas que representam as categorias resultantes dos componentes bioclimático, geofísico e edáfico (mapa_bioclimatico_pais.grd, mapa_bioclimatico_pais.gri, mapa_geofisico_pais.grd, mapa_geofisico_pais.gri, mapa_edafico_pais.grd e mapa_edafico_pais.gri). Todos estes mapas podem ser abertos em DIVA-GIS. Inicialmente DIVA-GIS abre os mapas como mostra a Fig. 17.

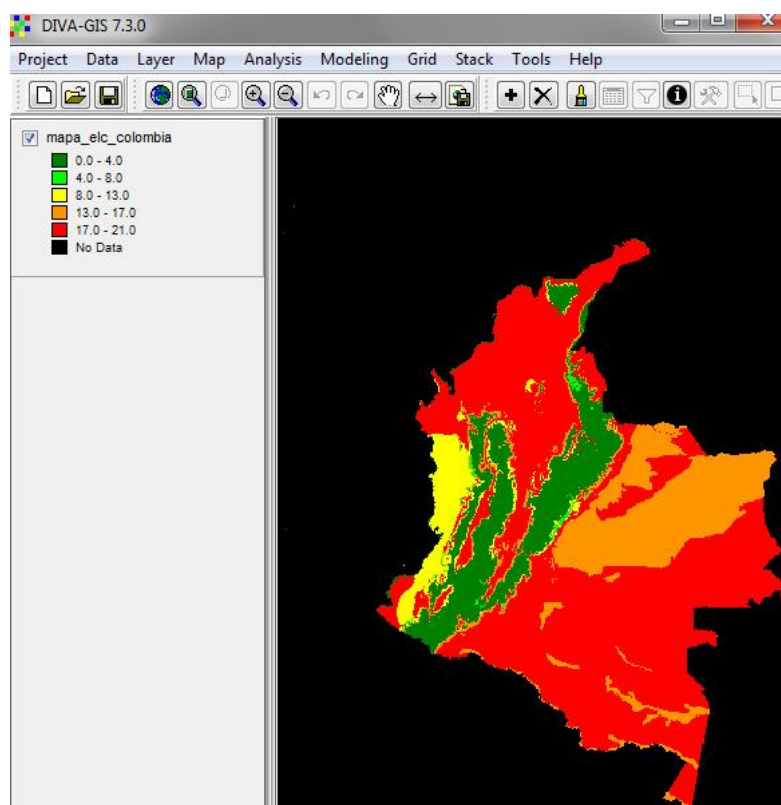


Figura 17. Exemplo de um mapa de Colômbia produzido por ELC mapas tal e como mostraria DIVA-GIS sem nenhum tipo de edição. As 20 categorias se mostram agrupadas em 5 classes.

Contudo, sua visualização pode-se alterar ao clicar duas vezes sobre a caixa cinza da parte esquerda que representa essa capa. Adicionando tantas filas até alcançar o número de categorias que contém o mapa, e aplicando uma gama de cores aleatórias, pode-se obter um mapa como o que mostra a Fig. 18. É conveniente que as cores sejam o mais diferentes possível, para que se identifiquem perfeitamente as categorias (cenários ecogeográficos) presentes no território.

NOTA: Lembre-se sempre que a categoria denominada como “0” (zero) não faz parte das categorias ecogeográficas do mapa resultado, mas é a denominação de todas as zonas para as quais existe informação de um ou dois componentes, mas não dos três. Por exemplo, por razões óbvias é

normal que a informação dos solos para zonas urbanas ou corpos de água seja nula, mas pode existir informação bioclimática e inclusive geofísica para estas áreas. Essas áreas serão codificadas como “0”.

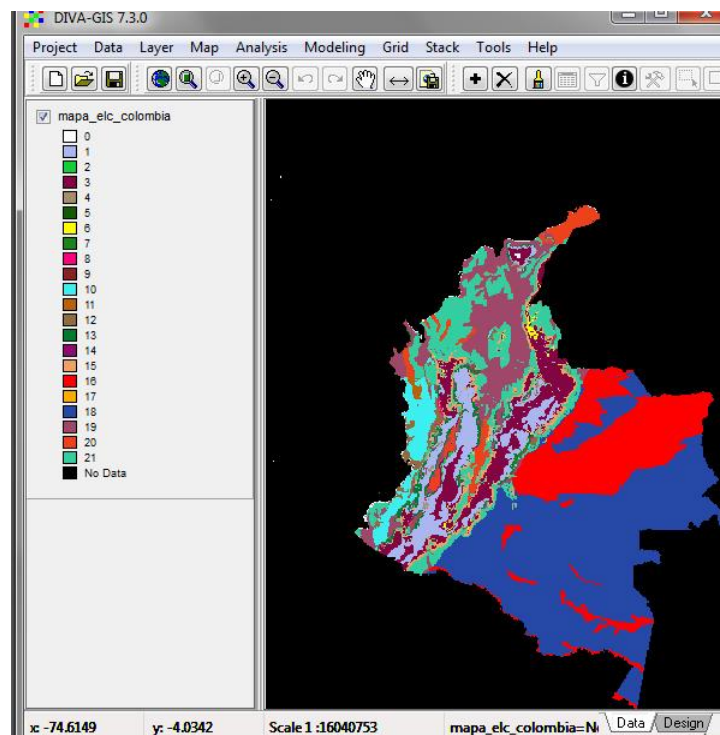


Figura 18. Exemplo de um mapa de Colômbia produzido por ELC mapas que mostra uma cor por cada uma de suas categorias. As propriedades do mapa tal e como abre DIVA-GIS foram alteradas para mostrar cada categoria com uma cor diferente.

Além dos mapas compatíveis com DIVA-GIS se produz um mapa compatível com Google Earth, “mapa_elc_pais.kml”. Sempre que se tenha instalado Google earth, se poderá abrir este mapa como uma capa sobre as imagens de Google earth ao clicar duas vezes sobre o arquivo. Este mapa não pode ser manipulado (mudança de cores) e sua qualidade gráfica não é ótima.

4.6.2 As tabelas

As tabelas produzidas por ELC mapas estão em formato de texto delimitado por tabulações e podem ser abertas em programas como Excel, OpenOffice, o R. Normalmente ao clicar com o botão direito do mouse, na opção “Abrir com” se oferecem alguns destes programas, se estão instalados.

4.6.2.1 “Tabla_ELC_celdas_pais.txt”. Esta tabela mostra os valores das variáveis selecionadas e os valores das categorias ELC (“ELC_CAT”), bioclimáticas, geofísicas e edáficas para cada centróide (fila) de cada célula que compõe o território do país objeto de estudo. Também inclui os valores de latitude e longitude de cada centróide.

4.6.2.2 “numero_categorias_pais.txt”. Contém uma simples recontagem das categorias ecogeográficas produzidas e representadas no mapa ELC resultante (columna “N_ELC_CAT”) e o número de categorias produzidas por componente.

4.6.2.3 “Estadist_ELC_pais.txt”, “Estadist_BIOCLIM_pais.txt”, “Estadist_EDAPH_pais.txt” y “Estadist_GEOPHYS_pais.txt”. Nestas tabelas encontram-se os estatísticos descritivos (média, valor mínimo, valor máximo e desvio padrão) para cada uma das variáveis originais envolvidas na criação do mapa ELC e para os mapas de cada um dos componentes (bioclimático, geofísico e edáfico) representados no mapa ELC. Estas tabelas se assemelham à tabela suplementária S2 apresentada à maneira da descrição das categorias do mapa ELC no artigo de Parra Quijano e colaboradores (2012 A).

4.7. Referências

Kaufman, L. y Rousseeuw, P.J. 1987, Clustering by means of Medoids, in Statistical Data Analysis Based on the L1-Norm and Related Methods. Y. Dodge (eds), North-Holland, 405–416.

Ketchen, D. J. y Shook, C. L. 1996. The application of cluster analysis in Strategic Management Research: An analysis and critique. Strategic Management Journal 17(6): 441–458.

Parra-Quijano, M. Iriondo, J.M., De la Cruz, M., Torres, M.E. 2011 A. Strategies for the development of core collections based on ecogeographical data. Crop Science 51:656-666

Parra-Quijano, M. Iriondo, J.M., Torres, M.E., De la Rosa, L. 2011 B. Evaluation and validation of ecogeographical core collections using phenotypic data. Crop Science 51:694-703

Parra-Quijano, M.; Draper, D.; Torres, E. e Iriondo, J.M. 2008. Ecogeographical representativeness in crop wild relative *ex situ* collections. p. 249-273. In Maxted, N.; Ford-Lloyd, B.V.; Kell, S.P.; Iriondo, J.M.; Dooloo, M.E. y Turok, J. (ed.) Crop wild relative conservation and use. CAB International, Wallingford.

Parra-Quijano, M. Iriondo, J.M., Torres, M.E. 2012 A. Ecogeographical land characterization maps as a tool for assessing plant adaptation and their implications in agrobiodiversity studies. Genetic Resources and Crop Evolution 59(2):205-217 DOI 10.1007/s10722-011-9676-7

Parra-Quijano, M. Iriondo, J.M., Torres, M.E. 2012 B. Improving representativeness of genebank collections through species distribution models, gap analysis and ecogeographical maps. Biodiversity and Conservation 21:79-96 DOI 10.1007/s10531-011-0167-0

Parra-Quijano, M. Iriondo, J.M., Frese, L., Torres, M.E. 2012 C. Spatial and ecogeographic approaches for selecting genetic reserves in Europe. En: N. Maxted, M.E. Dooloo, B.V. Ford-Lloyd, L. Frese, J. Iriondo y M.A.A. Pinheiro de Carvalho (ed.) Agrobiodiversity Conservation: securing the diversity of crop wild relatives and landraces. CABI, Wallingford, UK.

Rousseeuw, P.J. 1987. "Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis". Computational and Applied Mathematics 20: 53–65. doi:10.1016/0377-0427(87)90125-7.

Thormann, I. 2012. Applying FIGS to crop wild relatives and landraces in Europe. *Crop Wild Relative* 8 14:16.
<http://www.pgrsecure.org/publications>

Tohme, J., Jones, P., Beebe, S. y Iwanaga, M. 1995. The combined use of agroecological and characterisation data to establish the CIAT *Phaseolus vulgaris* core collection. p. 95-107. In Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L. y Morales, E.A.V. (eds.) Core collections of plant genetic resources. IPGRI, Rome.

5. Ferramenta ECOGEO

5.1. Caracterização ecogeográfica do germoplasma

Entende-se por caracterização ecogeográfica como a análise de toda informação ambiental do local onde cresce um indivíduo ou uma população vegetal, diretamente relacionada com o processo de adaptação ao entorno biótico ou abiótico. Em particular, as ferramentas CAPFITOGEN permitem analisar só o componente abiótico, classificado por três aspectos principais, os quais costumam considerar-se em estudos de adaptação de cultivos (Ceballos-Silva e López-Blanco, 2003) ou zoneamento agrícola (Williams et al., 2008):

- a) Bioclimático: Faz referência a fatores relacionados com temperatura e precipitação. Também inclui as relações entre temperatura e precipitação que se manejam através de índices.
- b) Geofísico: Agrupa os fatores topográficos e de destaque de maior importância, especialmente relacionados com a radiação solar.
- c) Edáfico: Fatores relacionados com as condições físicas e/ou químicas da fração do solo da qual as plantas dependem.

Portanto, caracterizar ecogeograficamente um conjunto de entradas implica assinalar a cada uma delas, informação de tipo bioclimático, geofísico e edáfico do local de sua coleta.

A informação ecogeográfica do local de coleta por si só mostra muitas características adaptativas do germoplasma, e em conjunto com outro tipo de caracterizações, como a fenotípica ou genotípica, pode resultar muito útil para explicar os padrões genéticos observados. Em alguns casos, como os que se dão em ausência de recursos econômicos suficientes para outro tipo de estudos, a caracterização ecogeográfica pode considerar-se uma alternativa válida, simples e barata para facilitar a utilização do germoplasma por parte de melhoradores que buscam parentais com certas características adaptativas nas coleções.

O insumo mais importante para realizar uma caracterização ecogeográfica são as coordenadas ou a descrição do local de coleta (do qual se possa extrair uma coordenada), que usualmente se registram nos descritores de passaporte no

momento da coleta. Usando tais coordenadas se irá assinalar a cada entrada os dados que descrevem as características ambientais mais importantes do local onde foi coletada. Isto implica que a qualidade dessas coordenadas é um aspecto crucial para uma designação correta de informação ecogeográfica, razão pela qual se sugere o uso da ferramenta GEOQUAL antes de realizar uma caracterização deste tipo.

Além das coordenadas como matéria prima, uma caracterização ecogeográfica requer informação ambiental que cubra toda a área de trabalho e de um programa especializado para a extração da informação correspondente a cada local de coleta. Os SIG são o tipo de programa que permite realizar este processo.

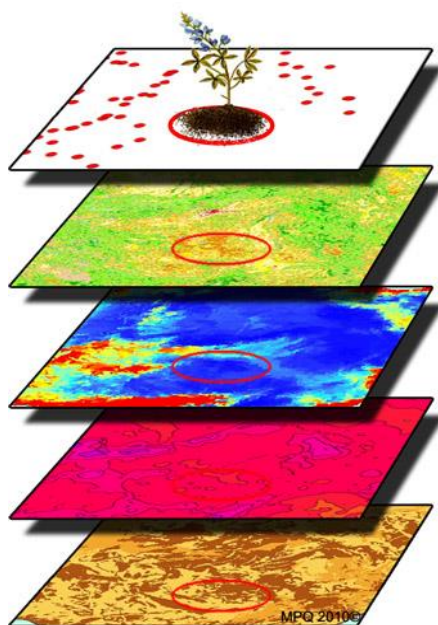


Figura 19. Ilustração do processo de extração de informação ecogeográfica que se realiza para um local de coleta através de SIG.

O produto de uma caracterização ecogeográfica é similar ao de qualquer outro tipo de caracterização, uma matriz de dados, onde as filas costumam corresponder às entradas e as colunas aos descritores. A partir desta matriz inicial é possível realizar análises multivariadas, como as que freqüentemente se realizam com outros tipos de caracterização, que permitem neste caso determinar relações de similaridade ambiental entre os diferentes locais de coleta. Uma análise fatorial, como por exemplo a de Componentes Principais (ACP), permitiria também conhecer as relações entre as diferentes variáveis originalmente introduzidas e criar variáveis sintéticas não correlacionadas que possam

descrever as afinidades ecogeográficas entre as entradas com um número reduzido de componentes.

É importante ressaltar que a caracterização ecogeográfica produz informação sobre os locais de coleta, não diretamente da natureza do germoplasma. Por isso, as análises de multivariados que funcionam sobre matrizes de distância ou dissimilaridade, neste caso, o que refletem é a afinidade ambiental e, indiretamente, a adaptação que existe entre os diferentes locais de coleta. Entradas de uma mesma espécie com padrões genotípicos ou fenotípicos diferentes podem ocorrer em cenários ambientais muito parecidos ou inclusive indiferenciáveis.

5.2. Características de ECOGEO

A ferramenta ECOGEO põe à disposição do usuário informação ecogeográfica de mais de cem variáveis (ver Anexos 12.1, 12.2 y 12.3), extraídas para uma lista de entradas que o usuário introduza na análise mediante o formato de dados do passaporte FAO/Bioversity 2012, com modificações menores, formato que se utiliza em várias das ferramentas CAPFITOGEN.

A ferramenta ECOGEO contém toda a informação necessária, o que permite que o usuário não tenha que descarregar nenhuma informação da internet (não requer conexão, como para as demais ferramentas CAPFITOGEN). A informação ou capas de variáveis ecogeográficas estão adaptadas e dispostas para funcionar com as programações R da ferramenta.

A área de trabalho é um segundo aspecto que define o usuário. Frequentemente a área de trabalho corresponde aos limites territoriais de um estado, tal e como vem definido na base de dados global sobre áreas administrativas (<http://www.gadm.org>). As variáveis ou camadas ecogeográficas estão recortadas segundo seus limites, de tal maneira que se escolhermos um determinado país e os dados de passaporte incluem coordenadas que indicam locais fora deste país, às entradas correspondentes a tais coordenadas não carregam nenhuma informação. Poderiam existir opções para fazer análises além das fronteiras de um país determinado, que é o que se considera uma região ou inclusive um continente. Se houver disponibilidade de regiões ou continentes dentro da ferramenta, (apareceriam nos listados do parâmetro “país”), o usuário pode usar

estas áreas de trabalho de maior cobertura, sabendo que seguramente o nível de resolução desta informação será de menor alcance (tamanhos de célula ao redor de 10x10 km).

Há um aspecto importante que introduz a ferramenta ECOGEO. Se trata da forma na qual se extrai a informação ecogeográfica para um local de coleta. Normalmente as extrações se realizam para o ponto que assinala a coordenada. Contudo, ocorrem duas situações nas quais a extração “pontual” não reflete as condições abióticas mais reais do local de coleta:

A) Quando se tem pouca informação sobre as coordenadas ou em geral estas não são de boa qualidade, segundo pode assinalar GEOQUAL ou outras metodologias. Inclusive em casos de espécies de distribuição costeira, onde apesar de um georreferenciamento relativamente apropriado, uma extração pontual poderia produzir muitos valores “NA” (informação não disponível) dado que os mapas/camadas ráster de informação ecogeográfica não se ajustam perfeitamente ao contorno das linhas costeiras.

B) Quando por diversas razões o local de coleta do germoplasma não corresponde exatamente ao local onde a planta cresce, mas este se encontra dentro de um perímetro relativamente conhecido (por exemplo quando se coleta germoplasma em mercados locais).

Nestos casos o usuário pode utilizar a extração do tipo “radial”. Nela, o usuário indica à ferramenta o raio da área circular ao redor do ponto que indica a coordenada, para a qual deseja extrair informação. Desta maneira, ECOGEO extrai dados ecogeográficos de todas as células que se encontram no interior da área circular, calcula seu valor médio e designa este valor à entrada, repetindo este processo para todas as variáveis ecogeográficas pelas quais se pretende caracterizar o germoplasma (ver Fig. 20). Além disso, ECOGEO descarta valores “NA” para o cálculo de médias e a subsequente designação de valores. Para indicar a ECOGEO que se deseja realizar uma extração “radial”, o usuário deve ativar o parâmetro buffy (seção 5.3.1.6) e à continuação introduzir o valor em metros do raio da área circular de extração no parâmetro tamp (seção 5.3.1.7).

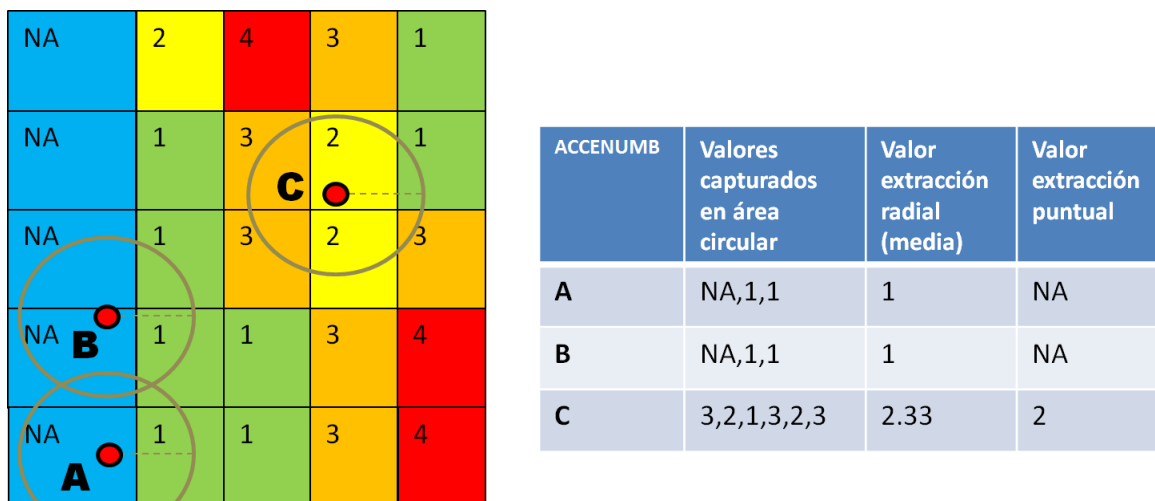


Figura 20. Diferenças entre os valores designados a partir de uma extração de tipo pontual e uma radial. As células de cor azul e valores NA representam corpos de água e os pontos de cor vermelha três locais de coleta (identificados com códigos ACCENUMB) localizados a partir de suas coordenadas.

Uma vez que o usuário preparou sua tabela de passaporte de acordo com o formato pré-estabelecido, indicou à ferramenta onde se encontra, e indicou a área de trabalho, o nível de resolução e a forma de extração que deseja, só resta definir as variáveis/camadas de interesse de cada aspecto (bioclimático, geofísico e edáfico) com as quais quer caracterizar os locais de coleta do germoplasma.

Com a definição destes parâmetros, a ferramenta ECOGEO buscará as variáveis/camadas de informação ecogeográfica de interesse, irá agrupá-las e extrairá a informação para cada coordenada a partir do grupo de camadas em um só passo. Com a informação extraída irá gerar uma tabela que será salva no endereço que o usuário defina no parâmetro “resultados”.

Finalmente, se o usuário estiver interessado em realizar uma análise de agrupamento ou uma Análise de Componentes Principais (ACP), pode indicar à ferramenta que execute tais análises e além disso pode assinalar que tipo de agrupamento deseja ou o número de componentes principais que deseja reter, respectivamente. A ferramenta ECOGEO produzirá gráficos (dendrogramas ou biplots) e tabelas (valores e vetores principais e pontuações para os componentes retidos) que se guardarão na pasta indicada no parâmetro “resultados”.

5.3. Utilização da ferramenta ECOGEO

Uma vez instaladas as ferramentas CAPFITOGEN e selecionada a ferramenta ECOGEO, será necessário definir uma série de parâmetros para que a programação R funcione corretamente.

5.3.1 Parâmetros iniciais definidos pelo usuário

5.3.1.1 Parâmetro: ruta

Explicação: Caminho é onde se encontram as ferramentas (documentos) CAPFITOGEN. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo: F:/CAPFITOGEN, C:/CAPFITOGEN, D:/MinhasFerramentas (MeusDocumentos)/CAPFITOGEN, etc.

5.3.1.2 Parâmetro: pais

Explicação: Selecionar o país para o qual se deseja construir o mapa ELC. Aparecerão vários países disponíveis somente se a ferramenta incluir informação para todos eles, se não, só aparecerá um determinado país escolhido.

5.3.1.3 Parâmetro: pasaporte

Explicação: Escrever o nome do arquivo que contém a tabela de passaporte em formato de texto, lembrando de incluir a extensão (.txt). Por exemplo, se o arquivo se chama "tabela" deve-se escrever: "tabela.txt". Recorde que este arquivo deverá ser guardado previamente na pasta "Passaporte", que faz parte do conjunto de pastas que compõe o diretório CAPFITOGEN.

5.3.1.4 Parâmetro: geoqual

Explicação: Selecionar esta opção se os dados do passaporte forem analisados pela ferramenta GEOQUAL e contêm 50 colunas (e não as 45 do modelo de passaporte que usam as ferramentas CAPFITOGEN). Desta forma, usar a tabela resultado de GEOQUAL v.2 denominada PasaporteOriginalEvaluadoGEOQUAL.txt como tabela de passaporte no ponto (item) anterior.

5.3.1.5 Parâmetro: totalqual

Explicação: Se sua tabela de passaporte provém de GEOQUAL e desejar possuir um mínimo de qualidade para que seus dados possam ser incluídos na análise, determinar o valor de TOTALQUAL100 a ser usado como limite. São admitidos valores de 0 (qualidade nula) a 100 (máxima qualidade).

5.3.1.6 Parâmetro: buffy

Explicação: Marcar esta opção se for desejável que a extração da informação ecogeográfica se faça sobre uma área ao redor do lugar de recolção. Deixar sem marcar esta opção levará a que a extração se realize somente para o ponto que indica as coordenadas do lugar de recolção.

5.3.1.7 Parâmetro: tamp

Explicação: Especificar o raio (em metros) de uma área circular ao redor do ponto que indica as coordenadas do lugar de recolção, de onde se extrairá a informação ecogeográfica. Os valores extraídos da área circular serão as médias para se obter um único valor e não serão consideradas células sem valor. Recomenda-se que este valor não seja inferior à distância de cada lado da célula do parâmetro resol1.

5.3.1.8 Parâmetro: resol1

Explicação: Selecionar o nível de resolução que se deseja utilizar para obtenção da informação ecogeográfica. Considerar que 1x1 km oferece melhor resolução, mas exige maior capacidade computacional em relação à 5x5 km, ainda que este não seja um critério tão limitante como para a ferramenta ELC mapas. As resoluções 10x10 e 20x20 km estão restritas aos países de grande extensão territorial, subcontinentes ou continentes. Ver o Anexo 12.5 sobre disponibilidade de resoluções em relação ao país ou região selecionada.

5.3.1.9 Parâmetro: bioclimsn

Explicação: Selecionar esta opção se for desejável a caracterização pelas variáveis de tipo bioclimático (temperatura, precipitação e índices associados).

5.3.1.10 Parâmetro: bioclimv

Explicação: Selecionar as variáveis bioclimáticas que se deseja incluir na caracterização ecogeográfica. O total das variáveis selecionáveis está detalhado no Anexo 12.1.

5.3.1.11 Parâmetro: edaphsn

Explicação: Selecionar esta opção quando for desejável caracterizar por variáveis do tipo edáfico (texturas, profundidade, pH, etc.).

5.3.1.12 Parâmetro: edaphv

Explicação: Selecionar as variáveis edáficas que se deseja incluir na caracterização ecogeográfica. O total das variáveis selecionáveis está detalhado no Anexo 12.2.

5.3.1.13 Parâmetro: geophyssn

Explicação: Selecionar esta opção quando for desejável caracterizar por variáveis do tipo geofísico (relativas ao relevo e radiação solar).

5.3.1.14 Parâmetro: geophysv

Explicação: Selecionar as variáveis geofísicas que se deseja incluir na caracterização ecogeográfica. O total das variáveis selecionáveis está detalhado no Anexo 12.3.

5.3.1.15 Parâmetro: latitud

Explicação: Incluir a latitude como variável de caracterização ecogeográfica?

5.3.1.16 Parâmetro: longitud

Explicação: Incluir a longitude como variável de caracterização ecogeográfica?

5.3.1.17 Parâmetro: ecogeoclus

Explicação: Selecionar esta opção quando for desejável realizar análise de agrupamentos (grupos) das entradas caracterizadas ecogeograficamente.

5.3.1.18 Parâmetro: ecogeoclustype

Explicação: Escolher o tipo de agrupamento hierárquico que se deseja utilizar para os agrupamentos (grupos) ecogeográficos: "single"= vizinho mais próximo, "complete" = vizinhança mais compactada, "ward" = método de mínima variância de Ward, "mcquitty" = método de McQuitty, "average" = similaridade (semelhança) média (UPGMA), "median" = similaridade (semelhança) da mediana, "centroid" = centróide geométrico, "flexible" = Beta flexível.

5.3.1.19 Parâmetro: ecogeopca

Explicação: Selecionar esta opção quando for desejável realizar análise de componentes principais para as entradas caracterizadas ecogeograficamente.

5.3.1.20 Parâmetro: ecogeopcaxe

Explicação: Número de componentes a considerar (sempre menor que o número de variáveis ecogeográficas) dentro da análise PCA.

5.4. Resultados de ECOGEO

No caminho e pasta que foram determinadas em “resultados” (parâmetro 5.3.1.21) se encontrará duas figuras e quatro tabelas.

5.4.1 As figuras

Se trata dos arquivos dendrograma_ecogeo.wmf e pca_ecogeo.wmf que correspondem à figuras vetoriais em formato metaarquivo do Windows. Estas figuras (um dendrograma e um biplot respectivamente) só irão gerar-se se foi solicitado à ferramenta realizar uma análise de agrupamentos (parâmetro 5.3.1.17) ou uma análise de componentes principais (parâmetro 5.3.1.19). Pode-se abrir e inclusive modificar no Microsoft Powerpoint ou programas de edição de imagens.

5.4.2 As tabelas

As quatro tabelas correspondem a dois tipos de informação.

5.4.2.1 Tabela de caracterização ecogeográfica das entradas: Se trata do arquivo TablaVarEcogeograficapais.txt. Corresponde à matriz inicial de

caracterização, com tantas filas como entradas analisadas e tantas colunas como descritores ecogeográficos.

5.4.2.2 Tabelas resultado de análises de componentes principais: Correspondem aos arquivos `ecogeographic_eigenvalues.txt` (tabela com os valores próprios), `ecogeographic_eigenvectors.txt` (tabela com os vetores próprios) e `ecogeographic_pcascors.txt` (tabela com as pontuações de cada entrada para os componentes principais retidos). Só irão gerar-se se foi solicitado à ferramenta realizar tal análise (parâmetro 5.3.1.19).

5.5. Referências

Ceballos-Silva, A. y López-Blanco, J. 2003. Evaluating biophysical variables to identify suitable areas for oat in Central Mexico: a multi-criteria and GIS approach. *Agriculture, Ecosystems and Environment* 95 (2003) 371–377.

Williams, C.L., Hargrove, W.W., Liebman, M. y James, D.E. 2008. Agro-ecoregionalization of Iowa using multivariate geographical clustering. *Agriculture, Ecosystems and Environment* 123 (2008) 161–174

6. Ferramenta Representa

6.1. Conceito de representatividade em coleções de germoplasma

Os aspectos sensíveis que podem pôr em risco uma conservação *ex situ* de recursos genéticos vegetais exitosa se podem dar em dois momentos particulares: ou no momento da coleta ou durante a conservação propriamente dita. Através da aplicação de técnicas apropriadas para o manejo do germoplasma, se pode reduzir o risco de perdas de entradas durante o período de conservação. Contudo, o germoplasma que se leva à conservação deve ser o reflexo da diversidade genética mais fiel possível das populações vegetais que ocorrem no campo. Durante a conservação no melhor dos casos e sem que mediem novas coletas, esse reflexo se manterá intacto. Esta situação põe em evidência a importância de realizar coletas de germoplasma que garantem capturar a maior diversidade genética possível. Quanto mais parecida a amostra da diversidade conservada *ex situ* em relação à diversidade genética total que ocorre na natureza, chamamos representatividade de uma coleção de germoplasma.

A representatividade de uma espécie dentro de uma coleção de germoplasma pode ser determinada a nível intra e interpopulacional. Para o caso de uma espécie cultivada, o equivalente seria níveis intra e intervarietais. São dois conceitos que, ao considerar-se de forma global a representatividade de uma coleção, são indissolúveis. Apesar disto, e por questões práticas relacionadas com a forma de conservação do germoplasma, ambos conceitos foram trabalhados a parte.

A representatividade intrapopulacional foi exaustivamente estudada como nos múltiplos trabalhos de Crossa e colaboradores (1994, 1997, 2011), o qual conduziu a estratégias de coletas específicas de acordo com a biologia reprodutiva da espécie, a distribuição espacial dos indivíduos e o tamanho da população. Basicamente o que se busca é calcular, segundo o caso, o número mínimo de indivíduos a coletar para garantir a captura da maior parte dos alelos presentes na população. Em contraste, a forma como deveria ser representada interpopulacionalmente uma espécie em uma coleção foi menos abordada.

Contudo, a partir do desenvolvimento do conceito de coleções núcleo ou nucleares, a representatividade interpopulacional de uma espécie em uma coleção ganha importância, dado que estas subcoleções só operam a este nível (Brown, 1989; Yonezawa et al., 1995).

Uma vez que o conceito de representatividade de uma coleção de germoplasma foi implantado na comunidade de científicos e curadores no âmbito dos recursos genéticos vegetais, o seguinte passo foi determinar a forma mais apropriada de calculá-la. Se o objetivo da conservação *ex situ* é capturar e manter a maior diversidade genética possível de uma espécie, a determinação ideal da representatividade seria em termos genéticos. Em consequência, a fórmula para determinar a representatividade genética (RG) em percentual seria:

$$RG = (NAC \cdot 100) / NAT$$

Onde NAT corresponde ao número total de alelos da soma de todos os loci estudados que apresenta a espécie objetivo, dentro do âmbito espacial (continente, país, região, etc.) da coleção a avaliar, e NAC é o número de alelos de tais loci desta espécie capturados por tal coleção. Esta determinação ideal da representatividade genética implica um impedimento de ordem prática. Conhecer o número total de alelos que pode conter uma espécie num território suficientemente grande como um país (espaço de trabalho normal de uma coleção de germoplasma de um Programa Nacional) ou inclusive a níveis muito inferiores a este, resulta uma tarefa na prática inviável para qualquer espécie, exceto para aquelas nas quais se sabe com certeza que estão compostas de pouquíssimas populações. Se considerarmos o contexto dos recursos genéticos vegetais para alimentação e agricultura, esta exceção é quase inexistente. Além disso, a tentativa de calcular a RG leva a que indiretamente se tenha representado 100% dos alelos, se a amostra de todas as populações implicar coleta de germoplasma. Em outras palavras, se calcular a RG de uma coleção de germoplasma implica coletar amostras e germoplasma de todas as populações que compõem a espécie dentro de um território de trabalho, então, independentemente do difícil que possa resultar esta tarefa, a máxima representatividade já seria alcançada sempre que na coleta tenha-se seguido critérios adequados de representatividade intrapopulacional.

Estas dificuldades práticas e logísticas levaram a se pensar em outras alternativas para determinar a representatividade de uma coleção. A representatividade ecogeográfica (RE) foi apresentada por Parra-Quijano e colaboradores (2008) no caso especial de coleções *ex situ* de plantas silvestres emparentadas com cultivadas (CWR em suas siglas em inglês). Nela se expõe a possibilidade de usar mapas de caracterização ecogeográfica do território (como os produzidos por ELC mapas) para conhecer quais condições ambientais presentes num marco espacial estariam representados em uma coleção de germoplasma.

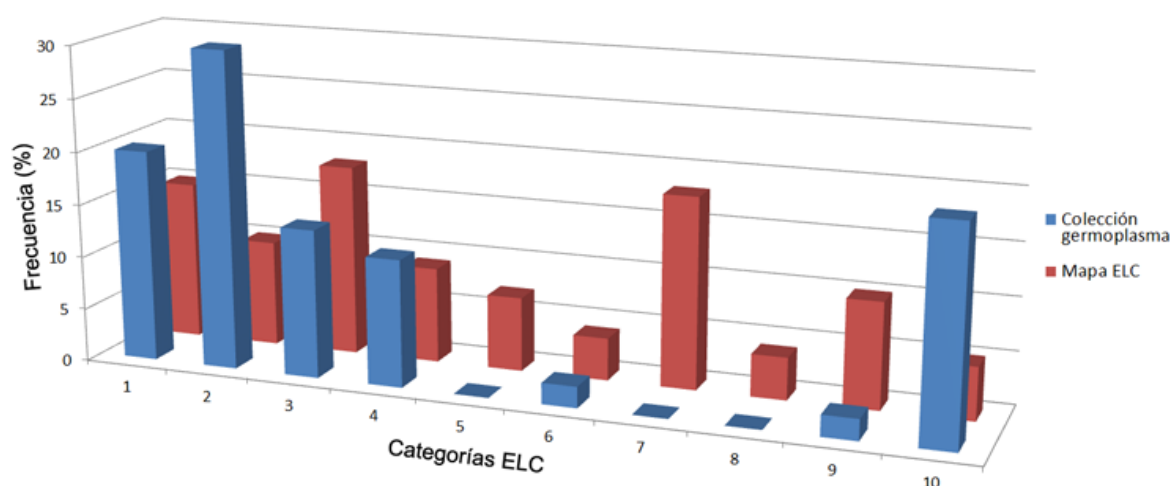


Figura 21. Comparação da representação de cada categoria ELC na coleção de germoplasma e a disponibilidade total de tais categorias no mapa ELC, medida em valores de frequências (em percentual).

Como exemplo de tal aplicação, a Fig. 21 mostra uma distribuição das frequências para cada categoria ELC de uma coleção de germoplasma em contraste com a disponibilidade das mesmas categorias no total do marco espacial. Este exemplo fictício serve para mostrar como a representatividade de uma coleção pode ou não (como é o caso) estar tendenciosa com respeito à abundância de unidades ambientais presentes no espaço de trabalho. O contraste de valores para as categorias 2 e 7 deixam antever que as duas distribuições não guardam semelhança e, possivelmente, uma prova de Qui quadrado determinaria uma associação não significativa das duas distribuições.

Contudo, a determinação mais exata da RE se consegue através de uma análise de faltantes ou vazios (mais conhecido por sua denominação em inglês como “gap analysis”). Para realizá-lo é necessário previamente compilar informações de outras fontes externas à coleção, como outras coleções de germoplasma ou

qualquer dado que indique a presença de populações da espécie objetivo (folhas de herbário, bases de dados botânicas, referências bibliográficas, etc.). Logo se compara a distribuição da frequência das coletas da coleção a ser avaliada com respeito às fontes externas e, então, se poderá visualizar claramente quais ambientes estão subrepresentados na coleção.

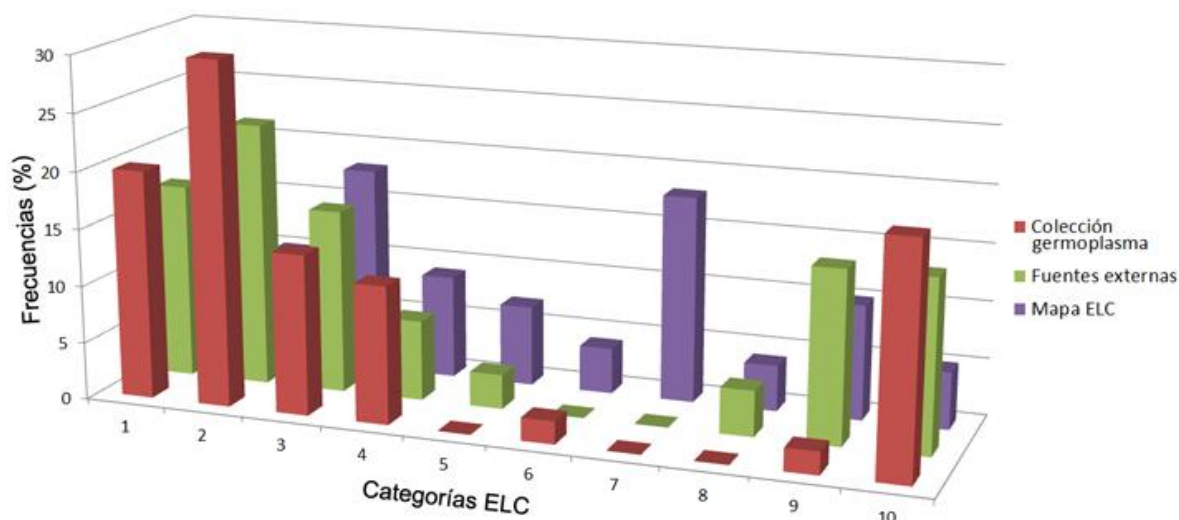


Figura 22. Comparação de distribuição das frequências de locais de coleta da coleção objetivo e de presenças de fontes externas sobre dez categorias ELC. Se inclui, também, a distribuição da frequência de cada categoria no total do mapa ELC.

A Fig. 22 ilustra o processo de comparação previamente mencionado. Usando os mesmos dados fictícios do exemplo da Fig. 21, neste gráfico de barras se inclui (em cor verde) a distribuição das frequências das categorias ELC para dados de presença provenientes de fontes externas. Neste caso é evidente a semelhança entre as distribuições coleção objetivo e fontes externas, ainda que algumas diferenças são especialmente interessantes. Para as categorias 5 e 8, as fontes externas indicam presença da espécie nessa unidade ambiental a qual não estaria representada na coleção, evidenciando a presença de faltantes ou vazios ecogeográficos. Estes vazios podem ser úteis para a planificação de novas coletas de germoplasma, pois se pode priorizar a visita destes ambientes já que se conhece onde ocorrem estas populações provenientes de fontes externas.

É importante deixar claro com respeito ao tratamento que se pode dar aos dados provenientes de fontes externas. Tomando as presenças provenientes de outro banco de germoplasma como fonte externa, se pode conhecer a representatividade da coleção objetivo globalmente, mas usar tais dados para

determinar locais prioritários para novas coletas pode levar a coletar duplicados inter-coleções.

6.3. Utilização da ferramenta Representa

Uma vez instaladas as ferramentas CAPFITOGEN e selecionada a ferramenta Representa, será necessário definir uma série de parâmetros para que a programação R funcione corretamente.

6.3.1 Parâmetros iniciais definidos pelo usuário

6.3.1.2 Parâmetro: ruta

Explicação: Caminho é onde se encontram as ferramentas (documentos) CAPFITOGEN. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo: F:/CAPFITOGEN, C:/CAPFITOGEN, D:/MinhasFerramentas (MeusDocumentos)/CAPFITOGEN, etc.

6.3.1.2 Parâmetro: internet

Explicação: Se desejar baixar informação sobre fontes externas da base de dados da internet (GBIF) e possuir acesso à rede, por favor, selecionar esta opção.

6.3.1.3 Parâmetro: pasaporte

Explicação: Escrever o nome do arquivo que contém a tabela do passaporte em formato de texto, lembrando de incluir a extensão (.txt). Por exemplo, se o arquivo se chama "tabela" deve-se escrever: "tabela.txt". Lembrar que este arquivo deve ser guardado previamente na pasta "Passaporte", que faz parte do conjunto de pastas que compõe o diretório CAPFITOGEN.

6.3.1.4 Parâmetro: geoqual

Explicação: Selecionar esta opção se os dados do passaporte forem analisados pela ferramenta GEOQUAL e contêm 50 colunas (e não as 45 do modelo de passaporte que usam as ferramentas CAPFITOGEN). Desta forma, usar a tabela resultado de GEOQUAL v.2 denominada PasaporteOriginalEvaluadoGEOQUAL.txt como tabela de passaporte no ponto (item) anterior.

6.3.1.5 Parâmetro: totalqual

Explicação: Se sua tabela de passaporte provém de GEOQUAL e desejar possuir um mínimo de qualidade para que seus dados possam ser incluídos na análise, determinar o valor de TOTALQUAL100 a ser usado como limite. São admitidos valores de 0 (qualidade nula) a 100 (máxima qualidade).

6.3.1.6 Parâmetro: fext

Explicação: Dispor de ocorrências provenientes de fontes externas (qualquer fonte de informação diferente à da coleção objeto de análise da representatividade) no formato solicitado?

6.3.1.7 Parâmetro: fuentex

Explicação: Indicar o nome do arquivo que contém as ocorrências de fontes externas no formato solicitado. Se o arquivo se chama "FontesExternas", neste campo deve aparecer "FontesExternas.txt" (porque esta tabela deve estar em formato de texto delimitado pelas tabulações). Lembrar que este arquivo deve encontrar-se na pasta chamada Passaporte.

6.3.1.8 Parâmetro: geoqualfe

Explicação: A tabela de ocorrências de fontes externas no formato solicitado dispõe, nas colunas indicadas, de informação sobre avaliação da qualidade do georreferenciamento (aplicação prévia do GEOQUAL)?

6.3.1.9 Parâmetro: totalqualfe

Explicação: Se sua tabela de ocorrências de fontes externas foi avaliada pelo GEOQUAL e desejar possuir um mínimo de qualidade para que seus dados possam ser incluídos na análise, determinar o valor de TOTALQUAL100 a ser usado como limite. São admitidos valores de 0 (qualidade nula) a 100 (máxima qualidade).

6.3.1.10 Parâmetro: duplibg

Explicação: Marcar esta opção se considerar como não existente (faltante) as ocorrências de fontes externas que provêm de outros bancos ou coleções de

germoplasma (Coluna TYPESOURCE com valor 40). Considerar que ao escolher esta opção, se poderia chegar a realizar coleções de populações que outras coleções já têm representadas, gerando duplicação intercoleção. Atenção: Marcar esta opção se todas as ocorrências de fontes externas que introduz provém de bancos ou coleções de germoplasma. Se não for assim será gerado erro.

6.3.1.11 Parâmetro: gbifFE

Explicação: É desejável baixar dados de ocorrência de fontes externas do portal da Global BiodiversityInformationFacility (GBIF)?. Esta opção requer conexão internet. Esta opção é incompatível com a inclusão de ocorrências de fontes externas informadas pelo usuário. No caso de escolher esta opção e aportar uma tabela com ocorrência de fontes externas, somente será considerada esta última.

6.3.1.12 Parâmetro: genero

Explicação: Escrever o nome do gênero da espécie objeto da presente análise. Este será o gênero para o qual se baixará informação do portal de GBIF. Lembrar que a primeira letra deverá ser maiúscula. Se não ativar o parâmetro gbifFE, não será necessário indicar nenhuma informação.

6.3.1.13 Parâmetro: especie

Explicação: Escrever o nome da espécie (somente o epíteto) objeto da presente análise. Este nome se unirá ao gênero para solicitar e baixar a informação desde o GBIF. Se desejar baixar informação para todo gênero, somente escreva aqui um *. Todo epíteto deve estar escrito em minúsculas. Se não ativar o parâmetro gbifFE, não será necessário indicar nenhuma informação.

6.3.1.14 Parâmetro: mapaelc

Explicação: Indicar o nome do arquivo que contém o mapa ELC (produto da aplicação da ferramenta ELCmapas), o qual deverá estar na pasta ELCmapas, que corresponde a uma das pastas que compõem o diretório CAPFITOGEN. Este mapa deve estar em formato DIVA-GIS, composto pelos dois arquivos

com extensões “.grd” e “.gri”, tal como produzido pelo ELCmapas. Nesta caixa de texto deve-se escrever o nome do arquivo com a extensão “.grd”. Assim, se o nome do mapa for "mapa_elc_brasil", você deverá escrever "mapa_elc_brasil.grd".

6.3.1.15 Parâmetro: statelc

Explicação: Indicar o nome do arquivo que contém a tabela de estatísticas descritivas do mapa ELC produzido pela ferramenta ELCmapas (esta ferramenta costuma nomear este arquivo como "Estadist_ELC_" mais o nome do país, correspondendo ao resultado 4.6.2.3). Igual que no mapa ELC, este arquivo deve estar situado na pasta ELCmapas. Da mesma forma, o nome deve ser escrito junto com a extensão e, neste caso, por ser uma tabela, a extensão será txt. Assim, se o arquivo se chamar "Estadist_ELC_brasil" deve-se escrever "Estadist_ELC_brasil.txt".

6.3.1.16 Parâmetro: distdup

Explicação: Determinar o valor de distância (em km) por baixo do qual se considerará que dois lugares de presença ou de recolção, na realidade, representam uma mesma população. O valor zero (por omissão) exclui da análise de representatividade entradas com mesmas coordenadas.

6.3.1.17 Parâmetro: resultados

Explicação: Introduzir o caminho da pasta onde deseja guardar os resultados das análises. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo C:/Resultados, D:/MinhasFerramentas(Meus Documentos)/Resultados, etc.

6.4. Resultados de Representa

No caminho e pasta que foi determinado em “resultados” (parâmetro 6.3.1.17) se encontrará até cinco mapas e até cinco tabelas.

6.4.1 Os mapas

Se trata de dois mapas de pontos de tipo vetorial (shapefiles) e três mapas raster (formato .grid) que podem ser visualizados diretamente em DIVA-GIS.

Se não se introduzem dados de fontes externas, a lista de mapas se reduziria a três (os descritos nas seções 6.4.1.1, 6.4.1.2 y 6.4.1.4).

6.4.1.1 “mapa_Class_ELC.grd”. Corresponde a um mapa que agrupa em quatro as categorias originais do mapa ELC. Estes grupos correspondem à divisão das categorias segundo sua frequência na totalidade do território. A frequência se divide a partir de quartos. O grupo indicado como 1 corresponde à frequência baixa (abaixo do quarto 0.25), o grupo 2 corresponde à frequência média-baixa (entre os quartos 0.25 e 0.5 ou média), o grupo 3 corresponde à frequência média-alta (entre os quartos 0.5 ou média e 0.75) e grupo 4 que corresponde à frequência alta (acima do quarto 0.75). Quando este mapa se abre em DIVA-GIS, apresentam-se cinco cores, por isso é importante mudar a visualização e que só se mostrem quatro cores, cada uma abarcando uma categoria que cubra cada valor de grupo (1 a 4). Esta mudança se ilustra na Fig. 23.

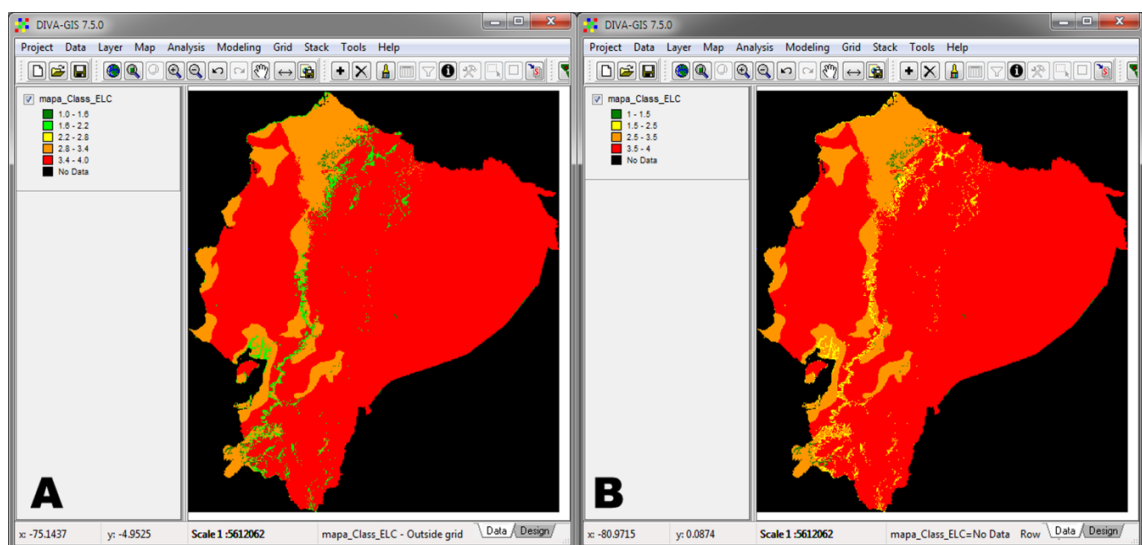


Figura 23. Configuração visual apropriada para mapas raster de Representa. A) Visualização mapa_Class_ELC.grd tal e como se abre em DIVA-GIS, B) visualização ajustada a quatro cores por cada grupo de frequência. Os cenários adaptativos menos frequentes (grupos baixa e média-baixa frequência) aparecem em cores verde e amarelo.

6.4.1.2 “mapa_Class_Sp.grd”. Corresponde a um mapa que agrupa em quatro as categorias originais do mapa ELC. Estes grupos correspondem à divisão das categorias ELC segundo a frequência em que ocorrem as entradas. A frequência se divide a partir de quartos. O grupo indicado como 1

corresponde à frequência baixa (abaixo do quarto 0.25), o grupo 2 corresponde à frequência média-baixa (entre os quartos 0.25 e 0.5 ou média), o grupo 3 corresponde à frequência média-alta (entre os quartos 0.5 ou média e 0.75) e grupo 4 que corresponde à frequência alta (acima do quarto 0.75).

6.4.1.3 “mapa_Tipo_faltante.grd”. Este mapa é outra reclassificação das categorias originais do mapa ELC. Este mapa só se produz quando o usuário introduz dados de fontes externas. Esta reclassificação corresponde aos critérios expostos na seguinte tabela:

Tabla 1. Classificação de categorias do mapa ELC de acordo com critérios de prioridade para futuras explorações

Classe	Diferença Fontes externas vs Banco de germoplasma (DIF) ¹	Classificação frequência por ocorrência da espécie ²	Classificação por frequência de categoria no mapa ELC ³
0	Não aplica	Não aplica	Não aplica
1	1	Baixa ou média-baixa	Baixa ou média-baixa
2	1	Baixa ou média-baixa	Média-alta ou alta
3	1	Média-alta ou alta	Baixa ou média-baixa
4	1	Média-alta ou alta	Média-alta ou alta
5	0.99-0.5	Baixa ou média-baixa	Baixa ou média-baixa
6	0.99-0.5	Baixa ou média-baixa	Média-alta ou alta
7	0.99-0.5	Média-alta ou alta	Baixa ou média-baixa
8	0.99-0.5	Média-alta ou alta	Média-alta ou alta
9	0.01-0.499	Baixa ou média-baixa	Baixa ou média-baixa
10	0.01-0.499	Baixa ou média-baixa	Média-alta ou alta
11	0.01-0.499	Média-alta ou alta	Baixa ou média-baixa
12	0.01-0.499	Média-alta ou alta	Média-alta ou alta
13	0 e NA	Não aplica	Não aplica

¹ Este valor é produto da comparação entre ocorrências de fontes externas e bancos/coleções de germoplasma em cada categoria usando a seguinte fórmula $DIF = (FE/BG)/FE$ sendo FE o número de ocorrências de fontes externas e BG as correspondentes a banco de germoplasma.

² Esta classificação é a mesma que se mostra no mapa 6.4.1.2

³ Esta classificação é a mesma que se mostra no mapa 6.4.1.1

Estas classes estão relacionadas com a prioridade de visita ou exploração que teria cada categoria ecogeográfica numa futura coleta. A classe 1 agrupa as categorias que teriam a máxima prioridade, enquanto que a classe 2 teria menos prioridade que a classe 1, e assim sucessivamente até a classe 13.

O mapa, ao ser aberto em DIVA-GIS não mostra as 13 classes com uma cor individual para cada classe, senão que agrupa os 13 valores em cinco cores.

Sua visualização correta se consegue em DIVA-GIS quando se acrescentam 8 cores mais e se ajustam as categorias de valores de cada cor (como nos mapas anteriores) ao valor de uma classe.

6.4.1.4 “Shapefile_Puntos_BG.shp”. Mapa vetorial (shafire) que representa os locais de coleta de germoplasma do banco ou a coleção de acordo com a avaliação de representatividade. A tabela que acompanha este mapa de pontos contém todos os campos do formato de passaporte FAO/Bioversity 2012.

6.4.1.5 “Shapefile_FE_class.shp”. Mapa vetorial (shafire) que representa as ocorrências de fontes externas. A tabela que acompanha este mapa de pontos apresenta os seguintes campos adicionais ao formato de dados de fontes externas:

FE_cat: Categoria do mapa ELC na qual ocorre a presença

FE_BG_dif: Valor DIF (ver Tabela 1) para a categoria ELC na qual ocorre a presença.

Class_Sp: Indica o quarto ao qual pertence a categoria na qual ocorre a presença de fonte externa, de acordo com frequência da espécie.

Class_ELC: Indica o quarto ao qual pertence a categoria na qual ocorre a presença de fonte externa, de acordo com a frequência da mesma categoria no mapa ELC.

Tipo_falt: Indica a que classe pertence a categoria na qual ocorre a presença de fonte externa de acordo com a classificação da Tabela 1.

6.4.2 As tabelas

Da mesma forma que com os mapas, a lista de tabelas pode ser reduzida de cinco a três, dependendo se o usuário introduz ou não dados de fontes externas.

6.4.2.1 “Tabla_Fuentes_Externas_clasificadas.txt”. Corresponde à mesma tabela acompanhante do shapefile da seção 6.4.1.5, contendo as mesmas variáveis.

6.4.2.2 “Tabla_Resultados_Representatividad.txt”. Nesta tabela se apresentam os resultados finais de avaliação da representatividade, se tenham ou não incluído dados de fontes externas. Com esta tabela é possível criar no Excel os gráficos de barras como os das Fig. 21 e 22. Finalmente nesta tabela se apresenta toda a informação requerida para o cálculo dos parâmetros da Tabela 1, incluindo o valor de classe com o que se definem as prioridades.

6.4.2.3 “Tabla_Resultados_X2.txt”. Esta tabela mostra os resultados da prova de Qui quadrado para determinar o grau de associação entre duas distribuições. Se foram introduzidos dados de fontes externas, esta tabela conterá os resultados de duas provas de Qui quadrado: distribuição banco/coleção (o BG) vs. fontes externas (FE) e banco/coleção vs. distribuição de freqüências totais das categorias no mapa ELC.

6.4.2.4 “TablaClasificacionCuartilesEspecie.txt” e “TablaClasificacionCuartilesMapaELC.txt”. Estas duas tabelas mostram os valores dos quartos 0.25, 0.5 (média) e 0.75 para a distribuição de freqüências da espécie e das categorias do mapa ELC.

6.5. Referências

Brown, A.H.D. 1989. The case for core collections. In : Brown, A.H.D., Frankel, O.H., Marshall, D.R. y Williams, J.T. (ed.) The use of plant genetic resources. Cambridge University Press, Cambridge, UK.

Crossa, J. y Vencovsky, R. 1994. Implications of the variance effective population size on the genetic conservation of monoecious species. Theoretical and Applied Genetics 89:936–942

Crossa, J. y Vencovsky, R. 1997. Variance effective population size for two-stage sampling of monoecious species. Crop Science 37:14–26

Crossa, J. y Vencovsky, R. 2011 Chapter 5: Basic sampling strategies: theory and practice. In: Guarino, L., Ramanatha Rao, V. y Goldberg, E. (ed.) Collecting Plant Genetic Diversity: Technical Guidelines – 2011 Update. Bioversity International. Available online (accessed 6 November 2013) http://cropgenebank.sgrp.cgiar.org/index.php?option=com_content&view=article&id=671

Parra-Quijano, M.; Draper, D.; Torres, E. e Iriondo, J.M. 2008. Ecogeographical representativeness in crop wild relative *ex situ* collections. p. 249-273. In Maxted, N.; Ford-Lloyd, B.V.; Kell, S.P.; Iriondo, J.M.; Dulloo, M.E. y Turok, J. (ed.) Crop wild relative conservation and use. CAB International, Wallingford.

Yonezawa, K.; Nomura, T. y Morishima, H. 1995. Sampling strategies for use in stratified germplasm collections. P. 35-53. *In* : Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L. y Morales, E.A.V. (ed.) Core collections of plant genetic resources. John Willey & sons, Chichester, UK.

7. Ferramenta DIV mapas

7.1. Representação espacial da diversidade local

Em 2012 foi publicado um estudo sobre a apresentação de padrões espaciais de diversidade genética a partir de marcadores neutros tipo microsatélite para o caso de *Annona cherimola* (van Zonneveld et al., 2012). Neste estudo se propõe uma forma diferente de mostrar como está distribuída a diversidade genotípica, baseado na estimativa de parâmetros próprios de genética de populações, só que antes de serem aplicados a todas as amostras ao mesmo tempo, neste caso a diversidade se estima a nível local com a determinação de vizinhanças ou áreas de influência. O resultado de juntar os resultados de cada vizinhança foi um mapa que mostra claramente onde se localizam os “pontos quentes” de diversidade. As aplicações desta metodologia na conservação *ex situ* e *in situ* dos recursos genéticos vegetais (variedades locais e espécies silvestres aparentadas com cultivadas) foram evidentes.

Esta não foi a primeira aproximação SIG ou de geoestatística na análise da diversidade genética, pois previamente se realizou interpolações de dados genéticos (Hoffman et al., 2003). Contudo, a metodologia de van Zonneveld e colaboradores representa uma aplicação muito prática e mais simples em sua análise e interpretação.

Posteriormente Thomas e colaboradores (2012) aplicaram essa mesma metodologia para 993 indivíduos caracterizados por microsatélites de cacau (*Theobroma cacao*), junto com outras análises, para identificar processos evolutivos nesta planta cultivada.

A partir da publicação destes desenvolvimentos foi possível entender, passo a passo, como opera o processo de obtenção de um mapa deste tipo. A metodologia é evidentemente repetível e só varia o parâmetro genético que se calcula entre as amostras que compõem uma vizinhança. Assim, se o parâmetro expressa que tão diferente geneticamente são as amostras de uma vizinhança, o mapa poderia denominarse como “mapa de diversidade”. A partir desta base, foi possível desenvolver a ferramenta “DIV mapas” e generalizar sua aplicação mais além de dados de caracterização genotípica.

É muito importante destacar que estes mapas mostram diversidade genotípica a nível intraespecífico, um aspecto que os diferencia notavelmente dos mapas de riqueza de espécies ou mapas de diversidade filogenética, os quais trabalham a nível interespecífico.

Mostrar a diversidade em forma de mapas tem múltiplas vantagens com relação às formas nas quais normalmente se apresentam estes resultados. Os mapas de diversidade, baseados no desenvolvimento original de van Zonneveld e colaboradores (2012) permitem identificar de maneira simples e rápida, as zonas ou regiões onde se concentra uma alta variabilidade. Um mapa destas características se converte numa poderosa ferramenta para a tomada de decisões em matéria de conservação *ex situ* e *in situ*.

7.1.1 Por que um mapa da diversidade ecogeográfica?

A diversidade ecogeográfica de um conjunto de acessos é uma medida das diferenças que ocorrem entre os cenários adaptativos de onde provêm tais acessos, ou em outras palavras, os locais de coleta. Se recorre ao termo “cenário adaptativo” antes que “ambiente”, porque para o cálculo da diversidade ecogeográfica só se considerariam as características ambientais de tipo abiótico de maior influência na distribuição e ocorrência da espécie objetivo, não todas as características ambientais disponíveis.

A diversidade ecogeográfica, como outro tipo de diversidade, se determina a partir de dados de caracterização do germoplasma. A caracterização ecogeográfica se realiza mediante a extração de informação para cada coordenada mediante a utilização de um SIG, o qual foi previamente carregado com camadas de informação ambiental.

A visualização da diversidade ecogeográfica em forma de mapa similar aos que desenvolveram van Zonneveld e colaboradores (2012), facilita a comparação entre zonas ou regiões com base à diferença dos cenários adaptativos onde ocorrem os pontos de coleta. As zonas ou regiões onde ocorrem maiores diferenças se podem traduzir diretamente em zonas onde cabe esperar germoplasma com adaptações mais divergentes e indiretamente uma possível ocorrência de uma maior diversidade genotípica ou fenotípica. Evidentemente, a determinação de zonas com maior diversidade genotípica ou fenotípica é ótima

quando se realiza através de dados de caracterização genotípica e fenotípica, respectivamente. Contudo, à falta destes, um mapa de diversidade ecogeográfica pode servir de solução intermediária enquanto se consegue caracterizar genotípica e/ou fenotipicamente as coletas. Em qualquer caso, o cenário ideal para a análise da diversidade segundo esta nova metodologia, é quando se podem obter mapas para os três tipos de caracterização, dado que seu contraste oferece uma visão biológica muito completa da situação dos recursos genéticos vegetais que ocorrem dentro de um marco de trabalho.

7.2. Procedimento de obtenção de mapas de diversidade de DIV mapas

DIV mapas é uma aplicação inspirada na aplicação de van Zonneveld e colaboradores (2012) para fruta do conde/ chirimoia (*Annona cherimola*), mas inclui algumas diferenças com respeito à metodologia original as quais são mais evidentes ao contrastar os dois processos. Nesta seção se mostrará passo a passo como DIV mapas consegue obter mapas de diversidade.

DIV mapas determina medidas de diversidade local, ou seja, entre os acessos coletados em uma zona com forma de grade com determinado tamanho e suas vizinhanças (zona de influência), usando como insumo informação ecogeográfica, fenotípica ou genotípica. Note-se que a partir deste ponto nos referiremos a entradas e não a amostras, dado que a ferramenta está pensada para ser utilizada no entorno de recursos genéticos vegetais, o qual não implica que não se possa usar em outros âmbitos biológicos. Como resultado, DIV mapas oferece uma ilustração gráfica que reflete os valores das medidas de diversidade distribuídas num espaço no modo do mapa, o qual facilita a visualização de zonas quentes de diversidade genética.

Para iniciar, é importante anotar que DIV mapas, assim como outras ferramentas incluídas no presente manual e muitas outras aplicações de SIG e ecogeografia em recursos genéticos vegetais, requer que cada entrada esteja georreferenciada e que, além disso, o esteja de maneira adequada. A seção 3 deste manual faz referência à ferramenta GEOQUAL que permite conhecer a qualidade do georreferenciamento do local de coleta do germoplasma. É aconselhável usar esta ferramenta antes de usar DIV mapas, de maneira que só entradas com suficiente qualidade em sua georreferenciação serão tidas em conta na obtenção de mapas de diversidade. Em todo caso, entradas sem coordenadas (campos

DECLATITUDE ou LATITUDE e DECLONGITUDE ou LONGITUDE) não poderão ser incluídas na análise que realiza DIV mapas.

O segundo ponto importante é que, se for necessário obter mapas de diversidade fenotípicos ou genotípicos, os dados de caracterização de cada tipo devem ser ajustados segundo o formato que se costuma fornecer na pasta “Formatos” (arquivos .xls do Excel). Se o tipo de mapa de diversidade requerido é do tipo ecogeográfico, DIV mapas inclui o mesmo processo de caracterização ecogeográfica de germoplasma que a ferramenta ECOGEO (capítulo 5), portanto para este aspecto não é necessário preparar tabelas ou matrizes de dados de caracterização senão simplesmente assinalar as variáveis ecogeográficas pelas quais se quer caracterizar as entradas.

DIV mapas aproveitará toda a informação de caracterização disponível e válida e, com base nela, realizará os mapas de diversidade para cada aspecto individualmente. Nesse sentido, as listas de entradas caracterizadas genotípica, fenotípica ou ecogeograficamente podem coincidir (o qual favorece a interpretação de resultados) ou não. O que é imprescindível é que os códigos de identificação das entradas nas tabelas de caracterização genotípica ou fenotípica, devem estar incluídos na tabela de passaporte FAO/Bioversity 2012 que contém a informação de georreferenciamento dos locais de coleta.

Uma vez esclarecidas estas condições, os seguintes pontos mostram como DIV mapas obtém os mapas de diversidade, independentemente dos dados de caracterização que use para isso.

7.2.1. Distribuição de locais de coleta e geração de retícula ou malha

A partir das coordenadas de cada local de coleta se gera um espaço de trabalho (x-min, y-min, x-max e y-max sendo x latitude e y longitude), sobre o qual se sobrepõe uma retícula ou conjunto de células de forma quadrada de um tamanho definido pelo usuário (ver Fig. 24). Além disso, se carrega uma camada com os centroides de cada célula da retícula (ver Fig. 25 parte A). Cada centroide tem um código de identificação.

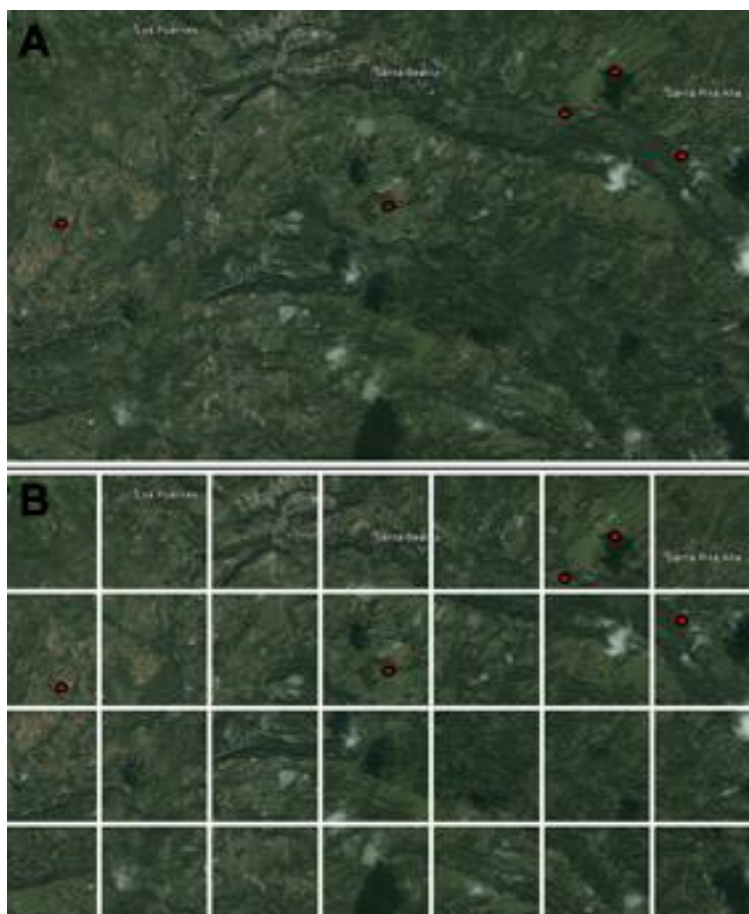


Figura 24. Primeiro passo. A) distribuição espacial dos locais de coleta, B) superposição de retícula de dimensões de célula (resolução) selecionadas pelo usuário.

7.2.2. Seleção de células com entradas e células vizinhas

Do total de células da retícula se selecionam aquelas onde ocorrem entradas. Além disso, o usuário determina uma área de influência, que assinala como o raio de uma área circular, a qual está relacionada com a biologia reprodutiva da espécie, o fluxo gênico e o manejo e a dispersão que pode gerar o homem, especialmente se se trata de uma forma cultivada. Esta área de influência se utilizará para determinar as células vizinhas, que são as células onde não ocorrem entradas mas que se encontram suficientemente perto das inicialmente selecionadas (células com entradas). Para ser célula vizinha, o centroide desta deve cair dentro da projeção da área circular de influência que se traça a partir do centroide de cada célula com entradas. O processo de seleção de células com entradas e de células vizinhas se mostra na Fig 25.

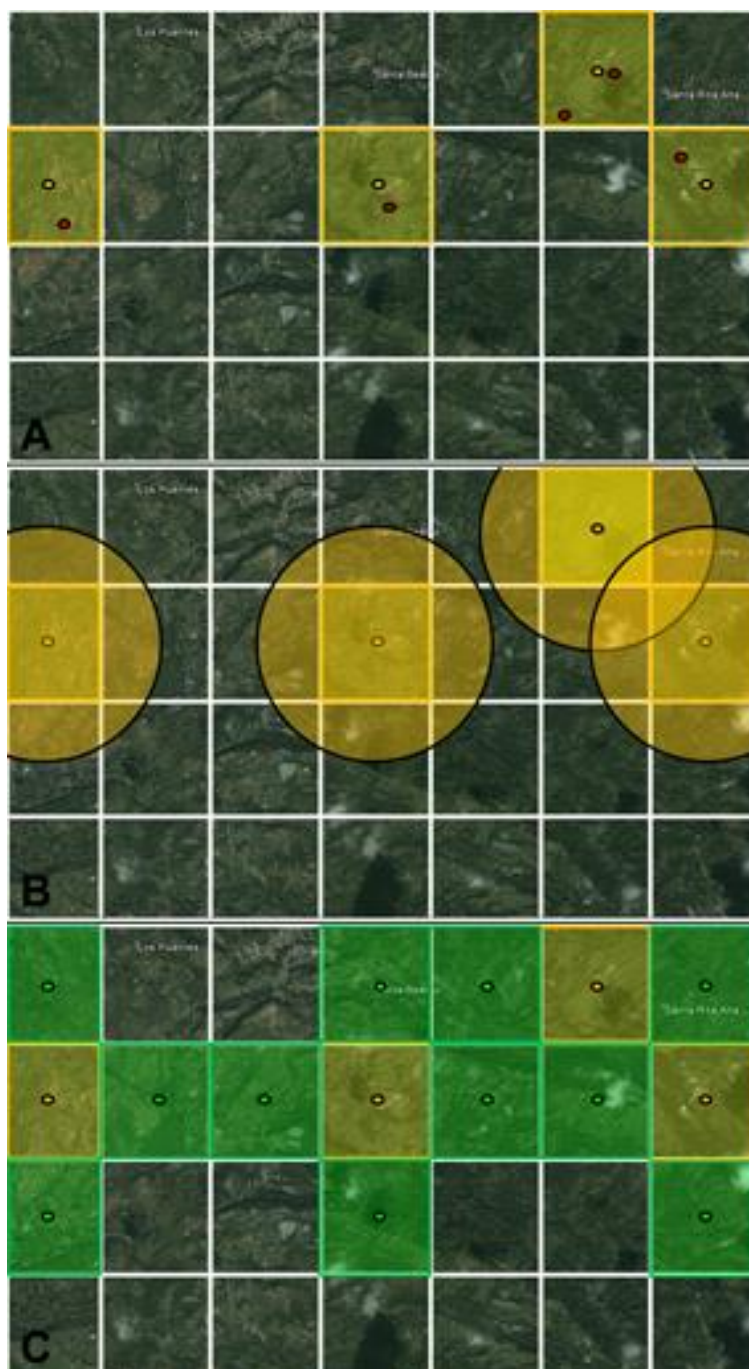


Figura 25. Segundo passo. A) Determinação de células com entradas e seus centroides, B) projeção das áreas de influência a partir dos centroides de células com entradas, C) determinação de células vizinhas.

7.2.3. Determinação de entradas associadas a célula com entradas e células vizinhas

A partir dos centroides das células com entradas e células vizinhas se projetam, outra vez, as áreas circulares de influência e à lista resultante de entradas que

caem em cada área se atribui o código de identificação de seu respectivo centroide (ver Fig. 26).

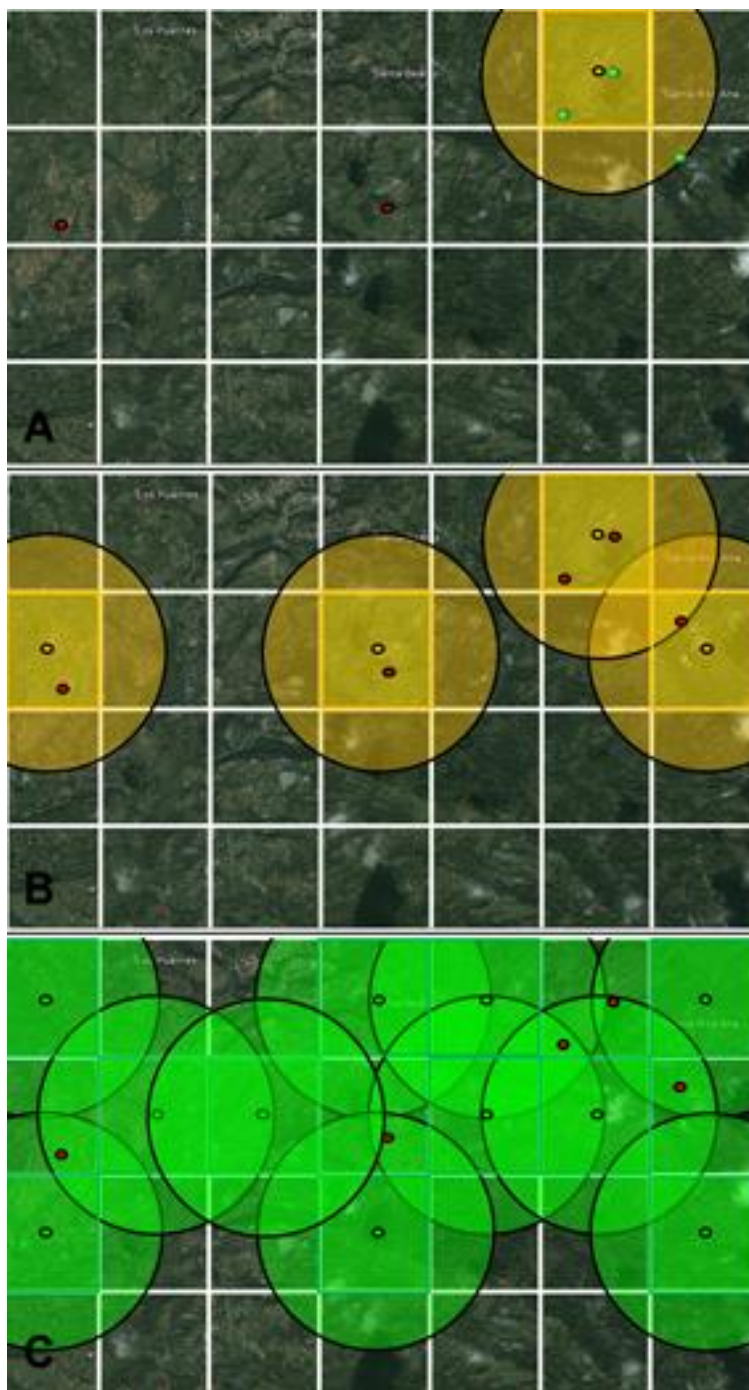


Figura 26. Terceiro passo. A) Determinação das entradas que ocorrem dentro da área de influência de uma célula em particular, B) determinação de entradas que ocorrem dentro de áreas de influência de células com entradas, C) determinação de entradas que ocorrem dentro de áreas de influência de células vizinhas.

7.2.4. Obtenção dos mapas finais de diversidade

A lista de entradas por células servirá para obter matrizes iniciais, com os dados de caracterização fenotípica, genotípica ou ecogeográfica (segundo os dados que tenha ingressado o usuário) como colunas, e as entradas de cada centroide identificadas por seu valor de ACCENUMB como filas. Desta forma se obterá tantas matrizes iniciais como células com entradas e células vizinhas forem determinadas. A cada matriz inicial se aplica estandardizações de dados se se trata de variáveis quantitativas e, posteriormente, um coeficiente de distância ou semelhança / não semelhança, o qual produzirá ao mesmo tempo, uma matriz diagonal de distância. A partir desta matriz, se calcula a distância média das entradas envolvidas e esse valor médio de distância se concederá a cada código de centroide e a sua respectiva célula. Isto permite a R produzir mapas de células raster que reflitam os valores atribuídos (ver Fig. 27).

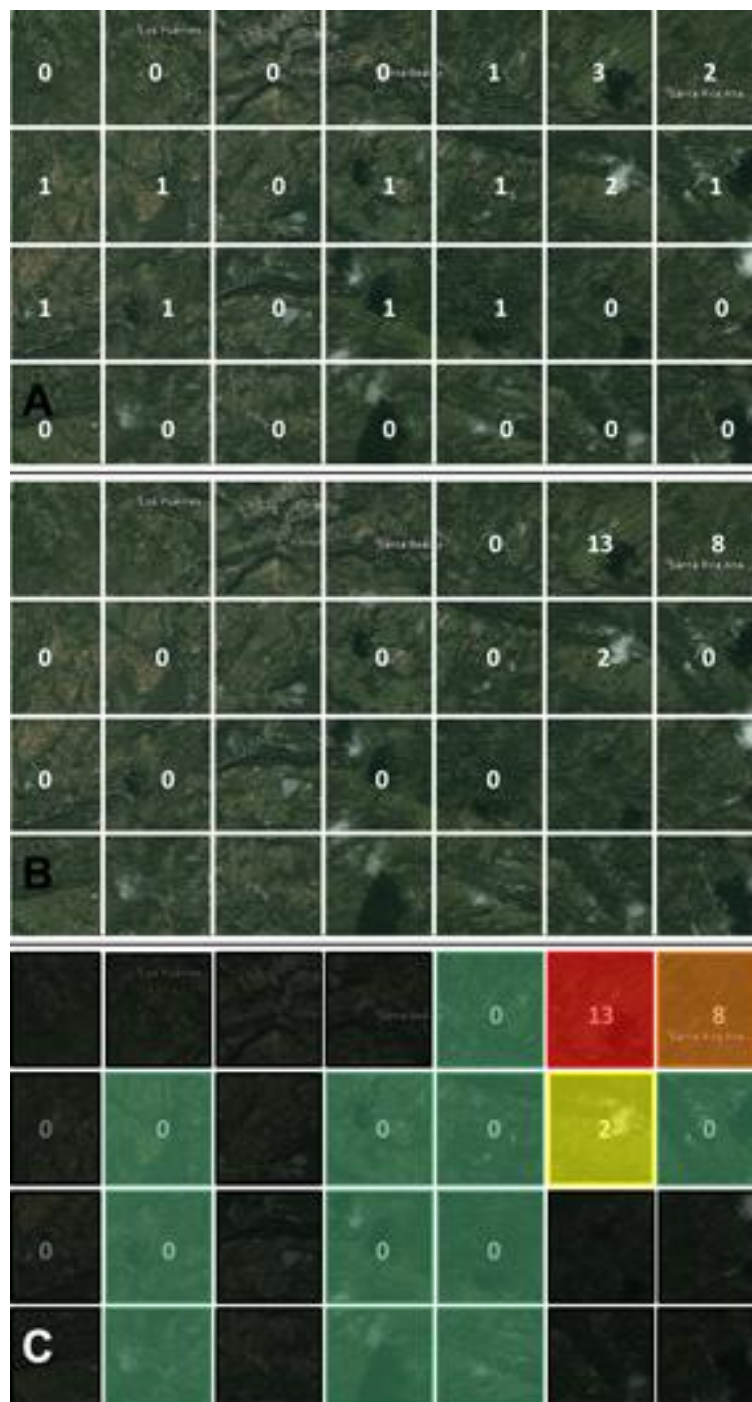


Figura 27. Quarto passo. A) Número de entradas analisadas por célula, B) valores atribuídos a células de distância genotípica, fenotípica ou ecogeográfica médias, C) Atribuição de graduação de cores de acordo com os valores médios de distância.

Para o caso de caracterizações genotípicas, adicionalmente à distância ou não semelhança média, se pode calcular outros parâmetros genéticos, como o da diversidade genética de Nei (1987) ou a proporção de marcadores polimórficos para cada grupo de entradas dentro de cada área de influência. Estes parâmetros são calculados por R a partir das matrizes iniciais de caracterização.

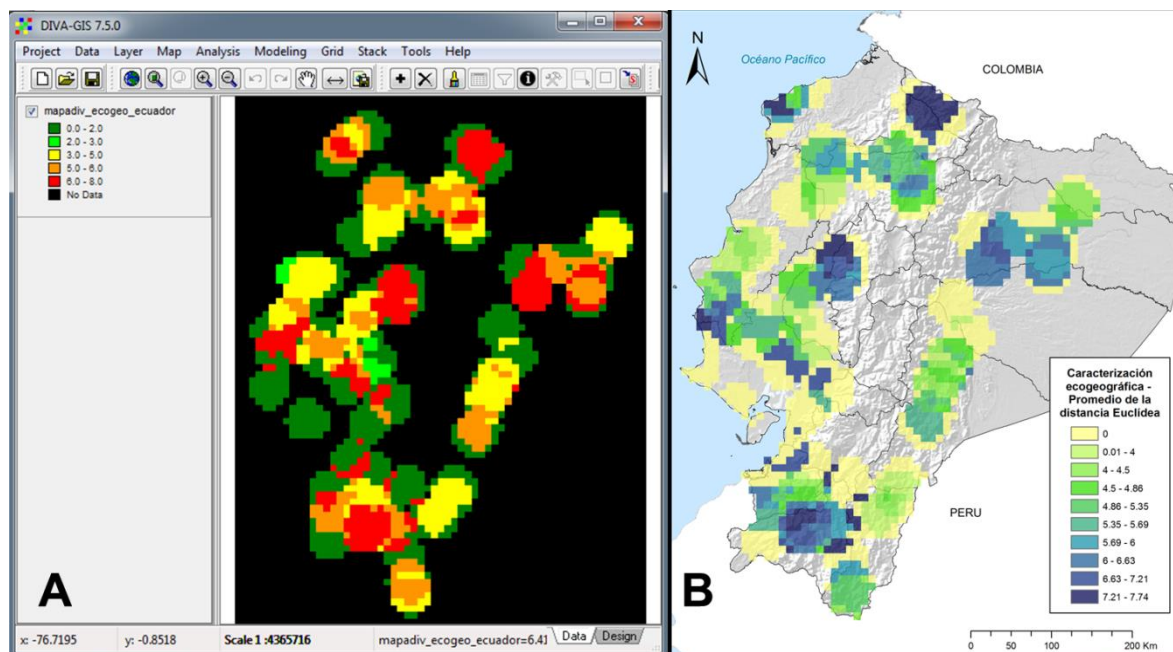


Figura 28. Visualização de mapas de diversidade. A) Mapa de diversidade ecogeográfica para a coleção equatoriana de amendoim (*Arachis hypogaea*) produto de DIV mapas aberto em DIVA-GIS, B) mesmo mapa em sua versão final.

Finalmente, ao visualizar-se em DIVA-GIS o arquivo ráster composto pelas células às quais se atribuiu valores dos parâmetros de diversidade, o que faz o programa informatizado é atribuir uma cor a cada célula de uma graduação de cores, permitindo observar rapidamente quais são os lugares onde se apresentam maiores níveis de diversidade, medida por valores médios de distância / não semelhança ou outros parâmetros genéticos (ver Fig. 28). Se se introduziu dados de caracterização de diferente tipo, se produzirão vários mapas, ocorrendo sempre a seguinte relação: um para o aspecto ecogeográfico, um para o aspecto fenotípico e um ou vários para o aspecto genotípico, dependendo se o usuário solicitou o cálculo de um ou vários parâmetros. Também se produz um mapa do número de entradas analisadas por célula como se mostra na Fig. 29, que corresponde ao que representa a Fig. 27 parte A. Este último mapa pode servir de apoio para determinar a existência de possíveis desvios ou inclinações na coleta ou na interpretação dos padrões que se possam encontrar nos mapas de diversidade.

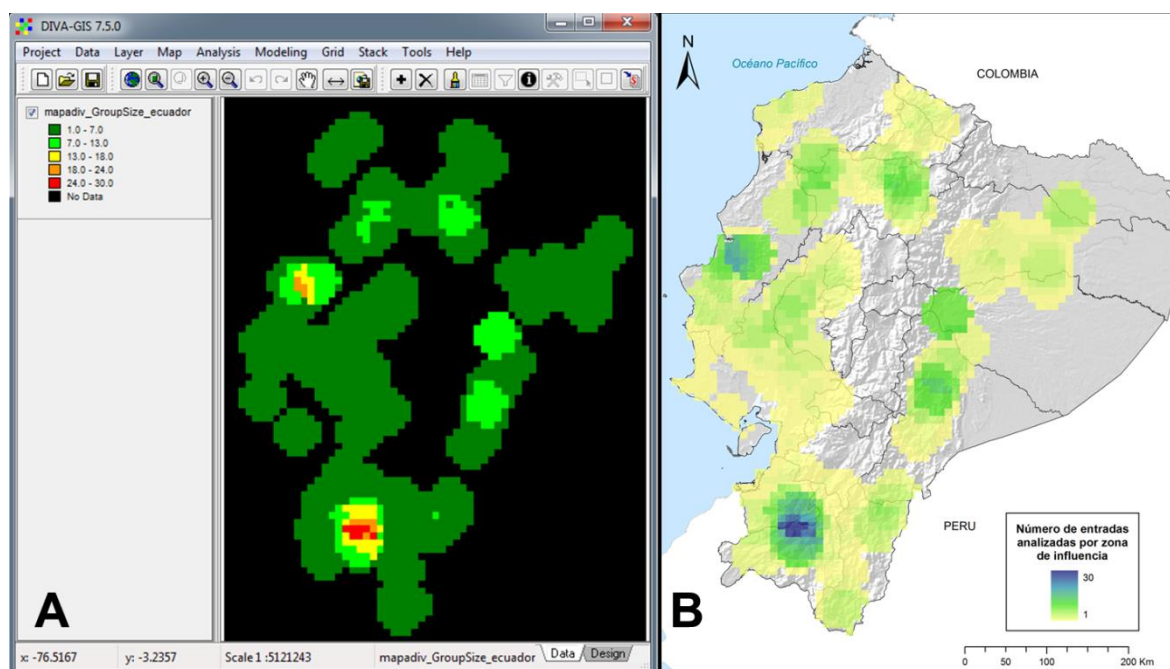


Figura 29. Visualização de mapa de número de entradas analisadas por célula. A) Mapa aberto em DIVA-GIS, B) mesmo mapa em sua versão final.

7.2.5. Uso de reamostragem para eliminação de possíveis desvios de coletas

van Zonneveld e colaboradores (2012) sugerem usar um método chamado de rarefação, o qual eliminaria o efeito de um possível desvio na coleta das amostras analisadas. Thomas e colaboradores (2013) sugerem um método de reamostragem. Este último método está incorporado em DIV mapas. Nele se usa como N (tamanho da amostra) a média do número de entradas analisadas por célula. Células com um número de entradas inferior à média são descartadas. Então o valor atribuído a cada célula corresponde à média dos valores médios de distância obtidos em cada processo de reamostragem. Dependendo do número de reamostragem que o usuário escolha e do número de células que componha o mapa de diversidade, o processo de obtenção pode tomar desde minutos até horas.

7.2.6. Outras análises

DIV mapas também permite realizar outros tipos de análises, particularmente quando foram introduzidos dados de caracterização de diferentes tipos. Desta maneira, DIV mapas pergunta ao usuário se deseja realizar análise de agrupamentos, análise de ordenação da mesma maneira que se realiza na ferramenta ECOGEO. Além disso, o usuário pode solicitar realizar comparações

matriciais de Mantel (1967) entre matrizes de distância para o total das entradas. De maneira automática, DIV mapas gera uma matriz de distâncias geográficas entre todos os locais de coleta e introduz esta matriz nas correlações matriciais pareadas.

7.3. Formatos de dados de entrada em DIV mapas

Para o correto funcionamento da ferramenta DIV mapas, é necessário introduzir informação de diversos tipos nos formatos indicados. Normalmente estes formatos se encontram na pasta “Formatos” dentro da estrutura de pastas e arquivos das ferramentas CAPFITOGEN. Dentro desta pasta encontraremos outra com o nome “Formatos DIVmapas” e dentro dela cinco arquivos Excel.

7.3.1. Modelo de dados de passaporte

Como em outras ferramentas CAPFITOGEN, os dados de passaporte devem ser ingressados em formato FAO/Bioversity 2012 com modificações menores (arquivo “ModeloDatosPasaporte_FAO_BIOVERSITY_2012.xls”). Como DIV mapas permite trabalhar com dados avaliados por sua qualidade de georreferenciação mediante a ferramenta GEOQUAL, além do modelo de dados de passaporte normal, se dispõe de um modelo que inclui campos adicionais para os valores de avaliação GEOQUAL (arquivo “ModeloDatosPasaporte_FAO_BIOVERSITY_2012_conDatosGEOQUAL.xlsx”). Contudo, a maneira mais fácil de usar dados de passaporte avaliados por GEOQUAL é introduzir diretamente a tabela de passaporte com dados de avaliação que produz GEOQUAL como resultado e que leva o nome de “PasaporteOriginalEvaluadoGEOQUAL.txt”. Lembre-se que esta tabela deve estar em formato de texto delimitado por tabulações, por isso deve ser exportada a partir do Excel neste formato, e deve ser localizada na pasta “Pasaporte” da estrutura de pastas e arquivos das ferramentas CAPFITOGEN.

7.3.2. Modelo de datos fenotípicos

Ao visualizar o conteúdo do formato de dados fenotípicos (arquivo “ModeloDatosFenotipicos.xlsx”) aparecerá uma coluna em verde (de preenchimento obrigatório) denominada “ACCENUMB” que corresponde ao mesmo código de identificação ACCENUMB da tabela de dados do passaporte. A ordem em que se encontram os códigos é indiferente. Dado que nem sempre se conta com dados de caracterização fenotípica para todas as entradas presentes

nos dados de passaporte, o número de entradas presente na tabela de dados fenotípicos pode ser menor ao de dados do passaporte. A situação que não deve ocorrer é que na tabela de dados fenotípicos apareçam entradas ou códigos ACCENUMB que não apareçam na tabela de dados de passaporte. Este último geraria um erro de processamento.

O resto das colunas no formato aparecem com os nomes “D1”, “D2” e “D3”. Estes nomes representam os nomes dos descritores fenotípicos 1, 2 e 3. O formato só inclui três colunas para os descritores, mas teoricamente podem ser tantos descritores quanto o usuário disponha, estendendo a sequência de “D4” até onde seja necessário. Os nomes podem ser mudados (por exemplo “D1” por “NGRANOS”), para facilidade do usuário. Se se deseja mudar os nomes, deve-se ter em conta duas recomendações. A primeira é de que não existam espaços dentro do nome e que, dentro do possível, o nome esteja composto por menos de onze caracteres e que, especialmente, não se repita nenhum nome. Esta última condição pode gerar erro.

A codificação das variáveis fenotípicas possui certas condições. As variáveis, sejam elas quantitativas ou categóricas, devem estar expressadas numericamente. Para as variáveis categóricas, os estados que durante a caracterização tenham sido descritos mediante códigos que incluem caracteres alfabéticos ou de outros tipos, devem ser transformados a códigos exclusivamente numéricos, sem linhas, pontos, vírgulas ou espaços. Se existem dados faltantes, estes devem ser codificados com os caracteres “NA”.

Finalmente tenha em conta que a ferramenta DIV mapas só reconhece a informação em tabelas quando esta se encontra em formato de texto separado por tabulações. Em consequência, uma vez a tabela de dados fenotípicos no Excel esteja completa e de acordo com os requerimentos previamente destacados, se deve exportar em formato de texto separado por tabulações e este arquivo de texto deve localizar-se na pasta “Pasaporte” junto com o resto de tabelas de dados de caracterização e a tabela de dados do passaporte.

7.3.3. Modelo de tabela de natureza das variáveis fenotípicas

No caso de que disponha de dados de caracterização fenotípica e que deseje utilizá-la para gerar um mapa de diversidade com DIV mapas, além de aportar a

tabela de dados fenotípicos da seção 7.3.2, é imprescindível preencher a tabela que aparece com o nome de “ModeloTablaNaturalezaVariables.xlsx”, a qual indicará a natureza de cada variável ou descritor fenotípico incluído na tabela de dados fenotípicos. Este arquivo Excel contém duas planilhas de cálculo.

Na primeira (“Natvariables”) se encontra a tabela de natureza de variáveis fenotípicas, que só contém três colunas. A primeira é “ID” e nela simplesmente se atribui um número consecutivo a cada variável (1, 2, 3,...), desta maneira cada fila nesta tabela corresponde a uma variável ou descritor fenotípico da tabela de dados fenotípicos. A segunda é “NOMVAR” e corresponde exatamente aos nomes atribuídos às variáveis ou descritores na tabela de dados fenotípicos. A terceira e última coluna é “NATVAR”, na qual se deve indicar a natureza da variável ou descritor correspondente. Ao situar o cursor sobre esta célula, se apresentará a lista de possíveis valores para esta coluna, a saber: binário simétrico, binário assimétrico, nominal, ordinal, quantitativo.

Finalmente na planilha de cálculo “Observaciones” se encontram algumas indicações e ajudas para preencher a tabela da planilha de cálculo “Natvariables”.

Ao finalizar, a tabela de natureza de variáveis deve ser exportada em formato de texto delimitado por tabulações e deve situar-se na pasta “Pasaporte”, de igual maneira que o resto de tabelas de ingresso de informação.

7.3.4. Modelo de dados genotípicos

Como se destacou previamente, DIV mapas oferece a possibilidade de fazer mapas de diversidade a partir de caracterizações genotípicas de germoplasma, analisando a informação proveniente de marcadores moleculares, como se estes fossem de tipo dominante. Isto implica que a tabela de dados genotípicos (arquivo Excel “ModeloDatosGenotipicos0_1.xlsx”) conterà variáveis de ausência/presença codificados como 0 e 1 respectivamente. Assim, a estrutura desta tabela é muito similar à de dados fenotípicos e se deve preencher de maneira similar, exceto porque todas as variáveis ou descritores na tabela de dados genotípicos correspondem a variáveis binárias assimétricas e se codificarão com valores 0 e 1.

Da mesma forma que com a informação fenotípica, tenha em conta que a ferramenta DIV mapas só reconhece a informação em tabelas quando esta se

encontra em formato de texto separado por tabulações. Em consequência, uma vez a tabela de dados genotípicos no Excel esteja completa e de acordo com os requerimentos assinalados, deve-se exportar em formato de texto separado por tabulações e este arquivo de texto deve localizar-se na pasta “Pasaporte” junto com o resto de tabelas de dados de caracterização e a tabela de dados de passaporte.

7.4. Utilização da ferramenta DIV mapas

Uma vez instalada as ferramentas CAPFITOGEN e selecionada a ferramenta DIV mapas, será necessário definir uma série de parâmetros para que a programação R funcione corretamente.

7.4.1 Parâmetros iniciais definidos pelo usuário

7.4.1.1 Parâmetro: ruta

Explicação: Caminho é onde se encontram as ferramentas (documentos) CAPFITOGEN. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo: F:/CAPFITOGEN, C:/CAPFITOGEN, D:/MinhasFerramentas (MeusDocumentos)/CAPFITOGEN, etc.

7.4.1.2 Parâmetro: pais

Explicação: Selecionar o país no qual foram recoletadas a maioria ou a totalidade das entradas que se deseja analisar. Se as entradas foram recoletadas em mais de um país, selecionar uma região, subcontinente ou continente (na sequência serão oferecidas estas opções).

7.4.1.3 Parâmetro: bootstrap

Explicação: Marcar esta opção se desejar calcular os valores para os mapas mediante bootstrapping (re-amostragem). Considerar que escolher esta opção reduzirá as áreas de análises nos mapas onde a densidade de lugares de coleção é menor.

7.4.1.4 Parâmetro: bootn

Explicação: Especificar o número de re-amostragens quando escolhida a opção de bootstrapping.

7.4.1.5 Parâmetro: replac

Explicação: Marcar esta opção quando desejar que as re-amostragens se realizem com substituição.

7.4.1.6 Parâmetro: pasaporte

Explicação: Escrever o nome do arquivo que contém a tabela de passaporte em formato de texto, lembrando de incluir a extensão(.txt). Por exemplo, se o arquivo se chama 'tabela' deve escrever: 'tabela.txt'. Lembrar que este arquivo deve ser guardado previamente na pasta 'Passaporte' que faz parte do conjunto de pastas que compõe a ferramenta DIVmapas.

7.4.1.7 Parâmetro: geoqual

Explicação: Selecionar esta opção quando os dados de passaporte foram analisados pela ferramenta GEOQUAL e contém, portanto, 50 colunas (e não as 45 do modelo de passaporte que usam as ferramentas CAPFITOGEN). Desta forma, usar a tabela resultado de GEOQUAL v.1.2, denominada 'PasaporteOriginalEvaluadoGEOQUAL.txt' como tabela de passaporte no ponto (item) anterior.

7.4.1.8 Parâmetro: totalqual

Explicação: Se sua tabela de passaporte é oriunda de GEOQUAL e é desejável considerar uma qualidade mínima para que seus dados sejam incluídos na análise, determinar o valor de TOTALQUAL100 a ser usado como limite. São admitidos valores de 0 (qualidade nula) a 100 (máxima qualidade).

7.4.1.9 Parâmetro: buffy

Explicação: Marcar esta opção se for desejável que a extração da informação ecogeográfica se faça sobre uma área ao redor do lugar de coleta. Deixar sem marcar esta opção levará a que a extração se realize somente para o ponto que indica as coordenadas do lugar de coleta.

7.4.1.10 Parâmetro: tamp

Explicação: Especificar o raio (em metros) de uma área circular ao redor do ponto que indica as coordenadas do lugar de recolção, de onde será extraída a informação ecogeográfica. Dos valores extraídos da área circular será feita a média para obter um único valor e não se considerarão células sem valor. Recomenda-se que este valor não seja inferior à distância de cada lado da célula do parâmetro resol1.

7.4.1.11 Parâmetro: ecogeo

Explicação: Selecionar esta opção se for desejável obter um mapa de diversidade ecogeográfica.

7.4.1.12 Parâmetro: resol1

Explicação: Selecionar o nível de resolução que se deseja utilizar para extração da informação ecogeográfica. Notar que 1x1 km oferece maior resolução, mas exige maior capacidade computacional em relação à 5x5 km, ainda que este não seja um aspecto tão limitante como na ferramenta ELC mapas. As resoluções 10x10 e 20x20 km ficam restritas aos países de grande extensão territorial, subcontinentes ou continentes. Ver o Anexo 12.5 sobre disponibilidade de resoluções em relação ao país ou região selecionada.

7.4.1.13 Parâmetro: bioclimsn

Explicação: Selecionar esta opção quando for desejável usar variáveis de tipo bioclimático (temperatura, precipitação e índices associados) para a geração de mapas de diversidade ecogeográfica.

7.4.1.14 Parâmetro: bioclimv

Explicação: Selecionar as variáveis bioclimáticas que se deseja incluir para gerar o mapa de diversidade ecogeográfica. O total de variáveis selecionáveis está detalhado no Anexo 12.1.

7.4.1.15 Parâmetro: edaphsn

Explicação: Selecionar esta opção se desejável usar variáveis do tipo edáfico (texturas de solo, profundidade, pH, etc.) para a geração de mapas de diversidade ecogeográfica.

7.4.1.16 Parâmetro: edaphv

Explicação: Selecionar as variáveis edáficas que se deseja incluir para gerar o mapa de diversidade ecogeográfica. O total de variáveis selecionáveis está detalhado no Anexo 12.2.

7.4.1.17 Parâmetro: geophysn

Explicação: Selecionar esta opção quando for desejável usar variáveis de tipo geofísico (relativas ao relevo e radiação solar) para a geração de mapas de diversidade ecogeográfica.

7.4.1.18 Parâmetro: geophysv

Explicação: Selecionar as variáveis geofísicas que se deseja incluir para gerar o mapa de diversidade ecogeográfica. O total de variáveis selecionáveis está detalhado no Anexo 12.3.

7.4.1.19 Parâmetro: latitud

Explicação: Incluir a latitude em seu mapa de diversidade ecogeográfica?
Nota: incluindo latitude e longitude, o aspecto de proximidade geográfica poderá ter algum reflexo no mapa final, ainda que somente a nível local (dentro de cada zona de influência).

7.4.1.20 Parâmetro: longitud

Explicação: Incluir a longitude em seu mapa de diversidade ecogeográfica?
Nota: incluindo latitude e longitude, o aspecto de proximidade geográfica poderá ter algum reflexo no mapa final, ainda que somente a nível local (dentro de cada zona de influência).

7.4.1.21 Parâmetro: phenotip

Explicação: Selecionar esta opção quando for desejável obter um mapa de diversidade fenotípica. Para isto deverá dispor de dados de caracterização ou avaliação fenotípica (morfologia, fenologia, produtividade, resistências, etc.) seguindo o formato indicado. Lembrar de incluir o nome da extensão. Desta forma, se por exemplo, a tabela se chamar 'fenotipos', neste espaço deverá se

escrever 'fenotipos.txt'. Lembrar que esta tabela deverá encontrar-se na pasta Passaporte, da estrutura de dados das ferramentas CAPFITOGEN.

7.4.1.22 Parâmetro: phenot

Explicação: Indicar o nome de arquivo de texto que contém os dados provenientes da caracterização fenotípica no formato indicado. Lembrar de incluir o nome da extensão. Desta forma, se por exemplo, a tabela se chamar 'fenotipos', neste espaço deverá se escrever 'fenotipos.txt'.

7.4.1.23 Parâmetro: phenotv

Explicação: Indicar o nome do arquivo de texto que contém a tabela que descreve a natureza de cada variável fenotípica no formato indicado. Lembrar de incluir o nome da extensão. Desta forma, se por exemplo, a tabela se chamar 'variablesfenotipo', neste espaço deverá se escrever 'variablesfenotipo.txt'. Esta tabela deverá conter todas as variáveis incluídas na tabela com os dados da caracterização (ponto ou item anterior).

7.4.1.24 Parâmetro: genotip

Explicação: Selecionar esta opção quando for desejável obter um mapa de diversidade genotípica. Para isto deverá dispor de dados de caracterização ou avaliação genotípica (como presença ou ausência de marcador como zeros e um), seguindo o formato indicado. Lembrar que esta tabela deverá encontrar-se na pasta Passaporte, da estrutura de dados das ferramentas CAPFITOGEN.

7.4.1.25 Parâmetro: genot

Explicação: Indicar o nome do arquivo de texto que contém os dados provenientes da caracterização genotípica no formato indicado. Lembrar de incluir o nome da extensão. Desta forma, se por exemplo, a tabela se chamar 'genotipos', neste espaço deverá se escrever 'genotipos.txt'.

7.4.1.26 Parâmetro: neigd

Explicação: Selecionar esta opção quando for desejável obter um mapa do índice médio de diversidade genética de Nei (1987), um mapa da proporção

média de marcadores polimórficos e um mapa de número de entradas analisadas por célula.

7.4.1.27 Parâmetro: csimilar

Explicação: Indicar o número correspondente do coeficiente de similaridade (semelhança) que se deseja utilizar para gerar o mapa de distância genotípica média. 1 = Índice de Jaccard (1901), 2 = SMC de Sokal&Michaner 1958, 3 = Sokal e Sneath (1963) (S5 de Gower e Legendre), 4 = Rogers&Tanimoto (1960), 5 = Dice (1945), 6 = coeficiente de Hamann, 7 = Ochiai (1957), 8 = Sokal e Sneath (1963) (S13 de Gower e Legendre), 9 = Phi de Pearson, 10 = S2 de Gower e Legendre. A distância (d) se obtém como $d=\sqrt{1-s}$, sendo s o coeficiente de similaridade (semelhança).

7.4.1.28 Parâmetro: rgrid

Explicação: Escolher o tamanho de célula (em km) que terá(ao) o/os mapas de diversidade que será(ão) gerado(s). Este parâmetro está restrito aos seguintes valores: 1, 5, 10, 50 e 100 km (se for escolhido outro valor se produzirá um erro).

7.4.1.29 Parâmetro: buffer

Explicação: Escolher o raio da área circular de influência ou vizinhança (em km). Esta área será criada a partir de cada centróide de célula do mapa com lugares de coleta e gerará grupos com as entradas, cujo lugar de coleta estará incluído. O valor de índices e distâncias médios de cada grupo será assinalado na célula daquele centróide em que se traçou a área de influência.

7.4.1.30 Parâmetro: ecogeoclus

Explicação: Selecionar esta opção quando for desejável realizar análise de agrupamentos (grupos) de todas as entradas para as quais se extraiu informação ecogeográfica.

7.4.1.31 Parâmetro: ecogeoclustype

Explicação: Escolher o tipo de agrupamento (grupo) hierárquico que se deseja utilizar para os agrupamentos (grupos) ecogeográficos: 'single'= vizinho mais próximo, 'complete' = vizinhança mais compactada, 'ward' = método de mínima variância de Ward, 'mcquitty' = método de McQuitty, 'average' = similaridade (semelhança) média (UPGMA), 'median' = similaridade (semelhança) da mediana, 'centroid' = centróide geométrico, 'flexible' = Beta flexível.

7.4.1.32 Parâmetro: ecogeopca

Explicação: Selecionar esta opção se for desejável realizar análise de componentes principais de todas as entradas para as quais se extraiu informação ecogeográfica.

7.4.1.33 Parâmetro: ecogeopcaxe

Explicação: Número de componentes a reter (sempre menor que o número de variáveis ecogeográficas) dentro da análise PCA.

7.4.1.34 Parâmetro: phenoclus

Explicação: Selecionar esta opção se for desejável realizar análise de agrupamentos (grupos) de todas as entradas para as quais se aportou informação fenotípica.

7.4.1.35 Parâmetro: phenoclustype

Explicação: Escolher o tipo de agrupamento (grupo) hierárquico que se deseja utilizar para os agrupamentos (grupos) fenotípicos: 'single'= vizinho mais próximo, 'complete' = vizinhança mais compactada, 'ward' = método de mínima variância de Ward, 'mcquitty' = método de McQuitty, 'average' = similaridade (semelhança) média (UPGMA), 'median' = similaridade (semelhança) da mediana, 'centroid' = centróide geométrico, 'flexible' = Beta flexível.

7.4.1.36 Parâmetro: phenopca

Explicação: Selecionar esta opção se for desejável realizar análise de componentes/coordenadas principais de todas as entradas para as quais foram aportadas informações fenotípicas.

7.4.1.37 Parâmetro: phenopcaxe

Explicação: Número de componentes/coordenadas a reter (sempre menor que o número de variáveis fenotípicas) dentro da análise PCA/PCoA.

7.4.1.38 Parâmetro: phenovarq

Explicação: Selecionar esta opção quando a totalidade das variáveis/descriptores fenotípicos corresponde a variáveis de tipo quantitativo.

7.4.1.39 Parâmetro: genoclus

Explicação: Selecionar esta opção se for desejável realizar análise de agrupamentos (grupos) de todas as entradas para as quais foi aportada informação genotípica.

7.4.1.40 Parâmetro: genoclustype

Explicação: Escolher o tipo de agrupamento (grupo) hierárquico que se deseja utilizar para os agrupamentos (grupos) genotípicos: 'single' = vizinho mais próximo, 'complete' = vizinhança mais compactada, 'ward' = método de mínima variância de Ward, 'mcquitty' = método de McQuitty, 'average' = similaridade (semelhança) média (UPGMA), 'median' = similaridade (semelhança) da mediana, 'centroid' = centróide geométrico, 'flexible' = Beta flexível.

7.4.1.41 Parâmetro: genopco

Explicação: Selecionar esta opção se for desejável realizar análise de coordenadas principais de todas as entradas para as quais foram aportadas informações genotípicas.

7.4.1.42 Parâmetro: genopcoaxe

Explicação: Número de coordenadas a reter (sempre menor que o número de variáveis genotípicas) dentro da análise PCoA.

7.4.1.43 Parâmetro: mantelt

Explicação: Indicar se for desejável analisar as correlações matriciais (Mantel, 1967) entre as possíveis combinações de fatores (ecogeográfico vs. fenotípico vs. genotípico). Realizar-se-ão todas as possíveis comparações, dependendo se foram introduzidos dados fenotípicos ou genotípicos ou se foi criada uma matriz ecogeográfica a partir de lugares de coleta. Será gerada e incluída uma matriz de distâncias geográficas nas comparações matriciais pareadas (par a par).

7.4.1.44 Parâmetro: mantelmeth

Explicação: Selecionar o tipo de correlação a usar na prova de Mantel.

7.4.1.45 Parâmetro: mantelper

Explicação: Ingressar o número de permutas que se deseja para realizar a prova de Mantel.

7.4.1.46 Parâmetro: resultados

Explicação: Introduzir o caminho da pasta onde deseja guardar os resultados das análises. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo C:/Resultados, D:/MinhasFerramentas(Meus Documentos)/Resultados, etc.

7.5. Resultados de DIV mapas

No caso de DIV mapas, a quantidade de gráficos, tabelas e mapas pode ser muito variável, dependendo dos dados que se introduzam, as opções que se indiquem e as análises que deseje realizar o usuário. Dado que a execução de DIV mapas pode chegar a produzir muitos resultados e que estes podem ser organizados de acordo ao tipo de dados e/ou análises de onde provêm, DIV mapas cria várias pastas dentro do caminho indicado no parâmetro resultados (ver seção 7.4.1.46). Desta maneira, os resultados são guardados nas correspondentes pastas temáticas que se explicarão nas seguintes seções, e só estarão fora delas o mapa de pontos correspondente aos locais de coleta em duas versões (“ShapefilePuntosPasaporte.shp” y “mapa_puntospas_google.kml”) e uma tabela (“Valor_mediana_para_bootstrapping.txt”), que só aparecerá quando se solicite

processos de reamostragem e que contém o valor da média utilizado como limiar em tal processo.

7.5.1 Pasta “ClassicMultivariateResults_pais”

Esta pasta contém gráficos (formato .wmf) e tabelas (.txt) produto de análises multivariadas (análises de agrupamentos e análises de componentes principais), tal e como se detalham nas seções 5.4.1 para as figuras e 5.4.2.2 para as tabelas. Dependendo dos dados introduzidos, aparecerá no nome dos arquivos os termos “ecogeographic” (provém da caracterização ecogeográfica), “genotypic” (provém da caracterização genotípica) e “phenotypic” (provém da caracterização fenotípica). Os resultados guardados nesta pasta correspondem aos produtos de análises de todas as entradas ao mesmo tempo, isto é, da maneira usual de análise, tal e como realiza a ferramenta ECOGEO.

7.5.2 Pasta “EcogeographicResults_pais”

Esta pasta aparecerá no caso de que se tenha solicitado a obtenção de um mapa de diversidade ecogeográfica (ver seção 7.4.1.11). Contém o mapa de diversidade em três formatos (imagen “mapadiv_ecogeo_pais.png”, Google earth “mapadiv_ecogeo_pais.kml”, e DIVA-GIS “mapadiv_ecogeo_pais.grd”). Nestes mapas aparecem representados com diferentes cores a média das distâncias ecogeográficas das entradas da área de influência de cada célula. A distância corresponde à euclidiana e tem possíveis valores desde 0 (para o caso em que tenha uma só entrada ou que todas as entradas tenham sido coletadas em ambientes idênticos) até o infinito.

Com respeito às tabelas, se pode encontrar as seguintes:

7.5.2.1 “tabla_estadisticas_mapadiv_ecogeo.txt”. Esta tabela mostra as estatísticas do mapa de diversidade ecogeográfica em termos de distância, ou seja, a média, desvio padrão e valores máximos e mínimos de distância, determinados no conjunto de células que compõem o mapa.

7.5.2.2 “TablaVarEcogeograficasecuador.txt”. Esta tabela contém os dados de caracterização ecogeográfica das entradas analisadas. Esta tabela se assemelha à que produz a ferramenta ECOGEO na seção 5.4.2.1.

7.5.2.3 “DistanciasMedias_sin_con_bootstrap_ECOGEO.txt”. Esta tabela só estará disponível quando se tenha solicitado realizar análises de reamostragem (seção 7.4.1.3). Nela aparecem os valores de distância média para cada célula (denominada aqui como “grupo”) sem reamostragem (coluna “WO_bootstrap”) e com reamostragem (“W_bootstrap”).

7.5.2.4 “TestBootstrapping.txt”. Esta tabela só estará disponível quando se tenha solicitado realizar análises de reamostragem (seção 7.4.1.3). Esta tabela mostra dois testes de associação entre os valores de distância média provenientes de processos com reamostragem e sem reamostragem (tabla 7.5.2.3).

7.5.3 Pasta “PhenotypicResults_pais”

Esta pasta aparecerá no caso de que se tenha solicitado a obtenção de um mapa de diversidade fenotípica (ver seção 7.4.1.11) e se tenha fornecido à tabela os dados correspondentes. Contém o mapa de diversidade em três formatos (imagen “mapadiv_phenot_pais.png”, Google earth “mapadiv_phenot_pais.kml”, e DIVA-GIS “mapadiv_phenot_pais.grd”). Nestes mapas aparecem representados com diferentes cores a média das distâncias fenotípicas das entradas da área de influência de cada célula. A distância corresponde a 1-coeficiente geral de similaridade de Gower (1971) e tem possíveis valores desde 0 (para o caso em que tenha uma só entrada ou que todas as entradas apresentem fenotipo idêntico) até 1 (máxima diferença).

Com respeito às tabelas, se pode encontrar as seguintes:

7.5.3.1 “tabla_estadisticas_mapadiv_phenot.txt”. Esta tabela mostra as estatísticas do mapa de diversidade fenotípica em termos de distância, ou seja a média, desvio padrão e valores máximos e mínimos de distância, determinados para todo o conjunto de células que compõem o mapa.

7.5.3.2 “TestBootstrapping.txt”. Esta tabela só estará disponível quando se tenha solicitado realizar análises de reamostragem (seção 7.4.1.3). Esta tabela mostra dois testes de associação entre os valores de distância média provenientes de processos com reamostragem e sem reamostragem.

7.5.4 Pasta “GenotypicResults_pais”

Esta pasta aparecerá no caso de que se tenha solicitado a obtenção de um mapa de diversidade genotípica (ver seção 7.4.1.11) e se tenha fornecido à tabela os dados correspondentes.

No interior da pasta se encontrarão os seguintes mapas:

7.5.4.1 “mapadiv_GenotDistance_pais”. Corresponde ao mapa de diversidade genotípica medida em distâncias médias em três formatos (imagen “.png”, Google earth “.kml”, y DIVA-GIS “.grd”). Nestes mapas aparecem representados com diferentes cores a média das distâncias genotípicas das entradas caracterizadas da área de influência de cada célula. A distância corresponde a 1-coeficiente de similaridade de Dice (1945) e tem possíveis valores desde 0 (para o caso em que tenha uma só entrada ou que todas as entradas apresentem fenótipo idêntico) até 1 (máxima diferença).

7.5.4.2 “mapadiv_GroupSize_pais”. Corresponde ao mapa de número de entradas analisadas por célula (imagen “.png”, Google earth.kml”, e DIVA-GIS “.grd”). Nestes mapas aparecem representados com diferentes cores o número das entradas de cada área de influência de cada célula.

7.5.4.3 “mapadiv_NeisGeneDiversity_pais”. Corresponde ao mapa de diversidade genotípica medida pelo índice de diversidade de Nei (1987) em três formatos (imagen “.png”, Google earth “.kml”, e DIVA-GIS “.grd”). Nestes mapas aparecem representados com diferentes cores o índice de diversidade antes mencionado, obtido das entradas caracterizadas da área de influência de cada célula.

7.5.4.4 “mapadiv_ProportionVariableMarkers_pais”. Corresponde ao mapa que mostra a proporção de marcadores polimórficos em três formatos (imagen “.png”, Google earth “.kml”, e DIVA-GIS “.grd”). Nestes mapas aparecem representados com diferentes cores a proporção de marcadores moleculares polimórficos obtida das entradas caracterizadas da área de influência de cada célula.

Com respeito às tabelas, se pode encontrar as seguintes:

7.5.4.5 “tabla_estadisticas_mapa_GenotDistance.txt”. Esta tabela mostra as estatísticas do mapa de diversidade genotípica (mapa 7.5.4.1) em termos de distância de Dice (1945), isto é, a média, desvio padrão e valores máximos e

mínimos de tal distância, determinados para todo o conjunto de células que compõem o mapa.

7.5.4.6 “tabla_estadisticas_mapa_NeiGeneDiversity.txt”. Esta tabela mostra as estatísticas do mapa de diversidade genotípica (mapa 7.5.4.3) em termos do índice de diversidade genética de Nei (1987), isto é, a média, desvio padrão e valores máximos e mínimos para este índice, determinados para todo o conjunto de células que compõem o mapa.

7.5.4.7 “NeiGeneDiversityMedias_sin_con_bootstrap.txt”. Tabela com os índices de diversidade genética de Nei (1987) obtidos para cada célula (aqui chamada “group”) sem reamostragem (coluna “WO_bootstrap”) e com reamostragem (“W_bootstrap”). Esta tabela só estará disponível quando se tenha solicitado realizar análises de reamostragem (seção 7.4.1.3).

7.5.4.8 “ProportPolymorphMarkersMedias_sin_con_bootstrap.txt”. Tabela com os valores de proporção de marcadores polimórficos obtidos para cada célula (aqui chamada “group”) sem reamostragem (coluna “WO_bootstrap”) e com reamostragem (“W_bootstrap”). Esta tabela só estará disponível quando se tenha solicitado realizar análises de reamostragem (seção 7.4.1.3).

7.5.4.9 “DistanciasMedias_sin_con_bootstrap_DICE.txt”. Nesta tabela aparecem os valores de distância média de Dice (1945) para cada célula (denominada aqui como “group”) sem reamostragem (coluna “WO_bootstrap”) e com reamostragem (“W_bootstrap”). Esta tabela só estará disponível quando se tenha solicitado realizar análises de reamostragem (seção 7.4.1.3).

7.5.4.10 “TestBootstrappingDICE.txt”. Esta tabela mostra dois testes de associação entre os valores de distância média de Dice (1945) provenientes de processos com reamostragem e sem reamostragem. Esta tabela só estará disponível quando se tenha solicitado realizar análises de reamostragem (seção 7.4.1.3).

7.5.4.11 “TestBootstrappingNei.txt”. Esta tabela mostra dois testes de associação entre os valores do índice de diversidade de Nei (1987) provenientes de processos com reamostragem e sem reamostragem. Esta tabela só estará disponível quando se tenha solicitado realizar análises de reamostragem (seção 7.4.1.3).

7.5.4.12 “TestBootstrappingPPM.txt”. Esta tabela mostra dois testes de associação entre os valores de proporção de marcadores polimórficos provenientes de processos com reamostragem e sem reamostragem. Esta tabela só estará disponível quando se tenha solicitado realizar análises de reamostragem (seção 7.4.1.3).

7.5.4 Pasta “MantelCorrelationResults_pais”

Nesta pasta se guardarão todas as tabelas com as matrizes de distâncias calculadas para todas as entradas ao mesmo tempo (“Matriz_distancia_”) e as tabelas com os resultados das correlações matriciais de Mantel (1967). O nome de cada tabela indica a comparação de informação realizada. Para as correlações onde intervêm dados genotípicos se usa a matriz de distância de Dice. Por exemplo, o arquivo “Mantel_genotypic_Vs_phenotypic.txt” contém os resultados da correlação matricial entre as distâncias genotípicas (Dice) e as distâncias fenotípicas (Gower). É importante notar que DIV mapas também calcula a matriz de distâncias geográficas (calculadas em graus decimais) para poder fazer comparações matriciais com respeito ao componente de distância geográfica.

7.6. Referências

Damme, P., Garcia, W., Tapia, C., Romero, J., Manuel Siguéñas, M. y Hormaza, J.I. 2012. Mapping Genetic Diversity of Cherimoya (*Annona cherimola* Mill.): Application of Spatial Analysis for Conservation and Use of Plant Genetic Resources. PLoS ONE 7(1): e29845. doi:10.1371/journal.pone.0029845

Dice, L.R. 1945. Measures of the Amount of Ecologic Association Between Species. Ecology 26:297–302.

FAO, BIOVERSITY. 2012. FAO/Bioversity multi-crop Passport descriptors V.2. Disponible en http://www.bioversityinternational.org/index.php?id=19&user_bioversitypublications_pi1%5BshowUid%5D=6901

Gower, J.C. 1971. A general coefficient of similarity and some of its properties. Biometrics 27: 857-74.

Hoffmann, M.H., Glaß, A.S., Tomiuk, J., Schmutz, H., Fritsch, R.M. y Bachmann, K. 2003. Analysis of molecular data of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae) with Geographical Information Systems (GIS). Molecular Ecology, 12: 1007–1019

Mantel, N. (1967) The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209-220.

Thomas, E., van Zonneveld, M., Loo, J., Hodgkin, T., Galluzzi, G., y van Etten, J. 2012. Present spatial diversity patterns of *Theobroma cacao* L. in the neotropics reflect genetic differentiation in pleistocene refugia followed by human-influenced dispersal. PLoS ONE 7(10): e47676. doi:10.1371/journal.pone.0047676

van Zonneveld M, Scheldeman X, Escribano P, Viruel MA, Van Damme P, et al. (2012) Mapping Genetic Diversity of Cherimoya (*Annona cherimola* Mill.): Application of Spatial Analysis for Conservation and Use of Plant Genetic Resources. PLoS ONE 7(1): e29845. doi:10.1371/journal.pone.0029845

8. Ferramenta ColNucleo

8.1. Conceito de coleção núcleo

Uma coleção núcleo ou nuclear (daqui em diante núcleo) é uma subcoleção, ou uma fração de uma coleção original, que se pode realizar por diversos motivos. O tamanho da coleção original é a maior determinante na hora de tomar a decisão de criar uma coleção núcleo. As coleções núcleo geralmente se consideram uma solução quando o tamanho das coleções originais se converte em um problema. Uma coleção de maior tamanho pode representar um problema na hora de multiplicar, caracterizar ou avaliar germoplasma, em particular, quando os recursos econômicos são limitados, selecionar materiais para programas de melhoramento, criar coleções ativas ou de trabalho, etc. Dependendo das condições de cada lugar, uma coleção de maior tamanho pode ser conformada por 500, 1000, 2000 ou mais entradas.

Uma coleção núcleo costuma ter um tamanho de 10% do total de entradas da coleção original, se bem que há estudos que situam o percentual ótimo acima ou abaixo deste valor (Parra Quijano et al., 2011a). A este percentual se denomina “intensidade de amostragem”.

É importante notar que de nenhuma maneira a determinação de uma coleção núcleo deve pôr em risco a conservação das entradas não selecionadas, as que se conhecem como “coleção de reserva”. Uma coleção núcleo pode facilitar a tomada de decisões em certas atividades de conservação e determinação de prioridades segundo condições de recursos insuficientes, o qual não exime a responsabilidade de conservar a coleção por completo. Por exemplo, quando é necessário realizar uma multiplicação de germoplasma, se conta com uma coleção núcleo e os recursos são limitados, se poderia pensar multiplicar, primeiramente, em primeiro termo as entradas da coleção núcleo e, com recursos adicionais em outro ciclo de multiplicação, o resto da coleção.

Independentemente dos motivos que levem à sua criação, a principal característica de uma coleção núcleo com respeito a outro tipo de subcoleções, é que esta deve ser representativa da diversidade genética contida na coleção

original. Isto implica que uma coleção núcleo deve conter entradas o mais diferentes entre si, de tal maneira que desde um ponto de vista genético, não se incluam entradas duplicadas ou altamente aparentadas (Brown, 1995). Consequentemente, para obter um subconjunto de entradas diferentes geneticamente, é imprescindível contar com informação sobre a composição genética da coleção, em outras palavras, dispôr de dados de caracterização.

Neste ponto, aparece um dos primeiros limitantes para obter coleções núcleo. Num cenário de recursos limitados, caracterizar uma coleção de mais de 1000 ou 2000 entradas pode resultar inviável. As caracterizações genotípicas e fenotípicas, geralmente, demandam desmedidos esforços econômicos e humanos que muitas instituições não podem assumir ou assumem de maneira parcial. Portanto, para superar este tipo de impedimentos, se pode utilizar outro tipo de dados de caracterização.

Em certos casos, quando foi requerida a criação de coleções núcleo e não se contavam com dados de caracterização para sua obtenção, a solução proposta foi o uso de dados de passaporte, particularmente, àqueles dados administrativos que descrevem a localidade do local da coleta (país, estado, província). Com isso, se pretendia assimilar diferentes unidades administrativas de coletas a diferentes ambientes e assim conseguir uma coleção núcleo representativa de todas as unidades administrativas e ambientes. Desta forma, se criaram diversas coleções núcleo administrativas em espécies como amendoim (Upadhyaya et al., 2003), guandul (Reddy et al., 2005), sésamo (Xiourong et al., 2000) ou sorgo (Grenier et al., 2001). Este tipo de coleções não assegurariam que a coleção núcleo estivesse formada pelas entradas mais diferentes em termos do ambiente onde foram coletadas, dado que as diferentes unidades administrativas são divisões antrópicas que não correspondem necessariamente a ambientes diferentes.

8.1.1 Estratégia de agrupamento.

O primeiro passo para a obtenção de uma coleção núcleo é a formação de grupos na coleção original por afinidade. Como se comentou previamente, para obter uma coleção núcleo é necessário contar com dados de caracterização ecogeográfica, genotípica, fenotípica ou no caso de coleções núcleo administrativas, dados de passaporte. Esta informação se usa para criar grupos

de entradas afins ou similares. A criação dos grupos geralmente é realizada através de métodos de classificação multivariada a partir dos dados de caracterização do germoplasma.

Numa variante para coleções núcleo de tipo ecogeográfico, se propõe não usar a caracterização do germoplasma que não seja do território onde este ocorre, através do uso de mapas ecogeográficos de caracterização do território (mapas ELC). Assim, as entradas são agrupadas de acordo com a categoria do mapa no qual ocorrem. Isto facilitaria a inclusão de novas entradas à coleção nuclear, dado que não seria necessário repetir a análise de agrupamentos, simplesmente conhecer o grupo (categoria ecogeográfica do mapa) ao qual pertence a nova entrada (Parra Quijano et al., 2011b).

8.1.2 Determinação de cotas por estratégias de afixação

Posteriormente, se determina o número de entradas que serão selecionadas de cada grupo de afinidade. Esse número ou cota é definida pela estratégia de afixação (*allocation strategy* em inglês) que o curador considere mais adequada. À medida que as coleções núcleo se popularizam, se propõem numerosas estratégias de afixação. Gradualmente, a complexidade e sofisticação de tais estratégias vai se incrementando. Contudo, alguns estudos comparativos mostram que as estratégias mais complexas não necessariamente são as que produzem coleções nucleares mais representativas (Parra Quijano et al., 2011b). As estratégias mais recorrentes, simples e testadas são as seguintes (Yonezawa et al., 1995):

- a) Aleatória (R): As entradas são selecionadas de maneira aleatória do total da coleção. Os grupos criados pela estratificação são ignorados.
- b) Constante (C): O mesmo número de entradas se seleciona de cada grupo, sem [levar](#) em conta o número de entradas que conformam o grupo.
- c) Proporcional (P): O número de entradas a selecionar de cada grupo é proporcional ao tamanho (número de entradas) de cada grupo.
- d) Logarítmico (L): O número de entradas a selecionar de cada grupo é proporcional ao logaritmo do tamanho (número de entradas) de cada grupo.

e) Dependente da diversidade (G): O número de entradas a selecionar de cada grupo é proporcional à diversidade contida em cada grupo. Esta estratégia requer acesso a dados de caracterização, adicionalmente aos agrupamentos provenientes da estratificação.

8.1.3 Informação sobre disponibilidade de entradas

Em boa parte dos estudos científicos sobre a criação de coleções núcleo, se realizam simulações para determinar qual é a estratégia de agrupamento e de afixação que produz a coleção núcleo mais representativa para cada caso, utilizando para isso o total da coleção. Contudo estas aproximações teóricas e simulações podem produzir coleções núcleo que na prática poderiam não se constituir devido ao fato de que as entradas selecionadas não estão disponíveis. A disponibilidade de uma entrada para ser parte de uma coleção núcleo, pode estar influenciada por vários fatores, entre eles o número de sementes disponível, se a entrada está só representada na coleção base ou se a entrada tem algum tipo de restrição em seu uso e distribuição. Por esta razão é importante que no caso da obtenção de uma coleção núcleo de carácter prático, se tenha em conta a informação de disponibilidade que possa fornecer o curador da coleção.

8.2. Coleções núcleo ecogeográficas

A caracterização ecogeográfica representa uma alternativa para a obtenção de coleções núcleo. Tendo em conta a relação entre fenótipo, genótipo e ambiente, uma coleção núcleo criada a partir de dados de caracterização ecogeográfica pode ser representativa não só das condições ambientais das populações de procedência das entradas contidas na coleção original, como também, de seus fenótipos e genótipos, sempre que a representatividade seja avaliada mediante características fenotípicas ou genotípicas relacionadas com a adaptação (Parra Quijano et al., 2011a).

O uso de dados de caracterização ecogeográfica do germoplasma na obtenção de coleções núcleo tem se documentado desde 1995, no caso do estabelecimento de uma coleção núcleo de *Phaseolus vulgaris* no Centro Internacional de Agricultura Tropical (CIAT) (Tohme et al., 1995). Dado que a ampla disponibilidade de SIG aplicados a recursos genéticos vegetais e de capas de informação ecogeográfica não foi possível até a década 2000-2010, as

coleções núcleo ecogeográficas não reaparecem no contexto científico internacional até 2008, com o caso de *Trifolium spumosum* (Ghamkhar et al., 2008).

Posteriormente, em dois estudos sobre diferentes tipos de coleções núcleo ecogeográficas, se determinou que a combinação mapa de caracterização ecogeográfica do território (mapa ELC) como estratégia de agrupamento e o tipo proporcional como estratégia de afixação, produziu coleções núcleo de alta representatividade ecogeográfica e fenotípica para coleções de *Lupinus* spp. e *Phaseolus vulgaris*, respectivamente (Parra Quijano et al., 2011a, 2011b). Em tais estudos, até 16 combinações diferentes de estratégias de agrupamento e afixação deram resultados similares ou inferiores em termos de representatividade ecogeográfica e fenotípica, que a combinação mapa ELC junto com afixação proporcional.

8.3. Obtenção de coleções núcleo ecogeográficas em ColNucleo

Seguindo as recomendações de alguns estudos científicos sobre coleções núcleo e representatividade, a ferramenta ColNucleo permite obter coleções núcleo ecogeográficas mediante a combinação de agrupamento por mapa ELC e três métodos de afixação (C, P y L). O mapa ELC deve ser gerado com a ferramenta “ELC mapas” (ver Capítulo 4).

A Fig. 30 mostra que, num primeiro passo, se extrai a categoria do mapa ELC correspondente ao local de coleta de cada entrada com coordenadas. Posteriormente, as entradas se agrupam de acordo com a categoria do mapa ELC designado. De acordo com a estratégia de afixação e a intensidade da amostragem que o usuário tenha selecionado, ColNucleo fixa as cotas ou número de entradas de cada grupo que conformarão em conjunto a coleção núcleo. Depois, ColNucleo determina se a cota pode ser cumprida com entradas que não correspondam a duplicados geográficos (não necessariamente genético) e que o curador tenha designado como “disponíveis” no caso de que o usuário tenha indicado a opção de utilizar dados de disponibilidade. As entradas não duplicado terão prioridade sobre entradas duplicado. Se a cota é de menor tamanho, que o número de entradas não duplicadas disponíveis, se selecionará aleatoriamente entre estas. Se a cota é de maior tamanho se selecionarão todas as entradas

disponíveis não duplicadas e o faltante se cobrirá com entradas duplicadas selecionadas aleatoriamente. Finalmente, as entradas selecionadas serão indicadas com o número 1 (um) numa nova coluna que se acrescentará à tabela de passaporte de entrada. No caso de haver utilizado dados de disponibilidade, as coleções núcleo obtidas podem ser incompletas ao não se dispor de suficientes entradas para representar uma ou várias categorias ELC. Por esta razão, ColNucleo gera uma tabela adicional na qual indica que acessos seriam necessárias estar disponíveis para que a coleção núcleo represente todas as categorias ELC de acordo com as cotas fixadas.

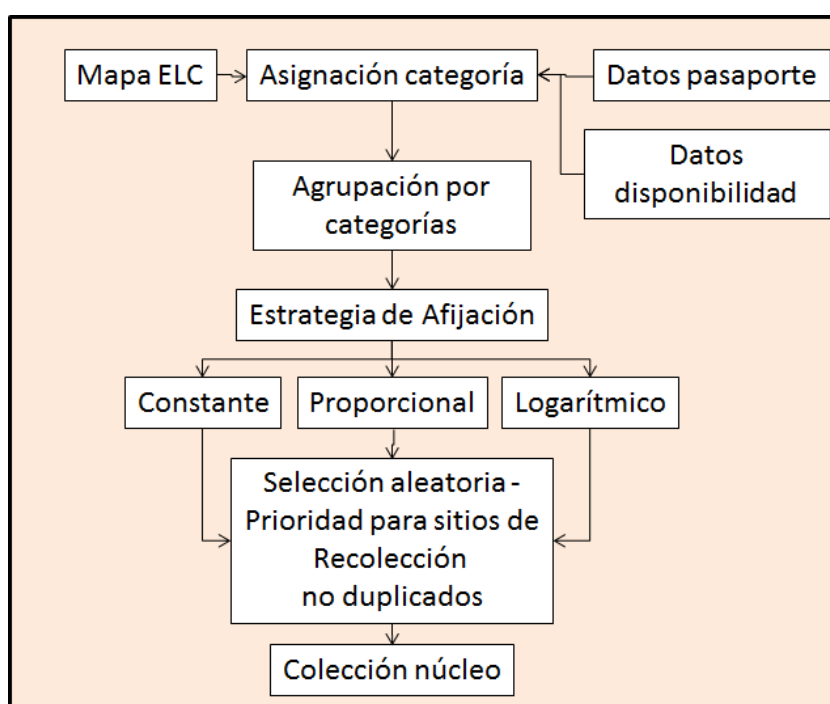


Figura 30. Ilustração do processo que segue a ferramenta ColNucleo para a obtenção de coleções núcleo ecogeográficas.

8.4. Formato de tabela de passaporte para ColNucleo

ColNucleo utiliza a tabela de passaporte FAO/Bioversity 2012 com modificações que utiliza as ferramentas GEOQUAL, Representa e ECOGEO, com a adição de um campo na parte direita, denominado “AVAILAB” que determina a disponibilidade de cada entrada. Desta maneira, na coluna AVAILAB se codificará com o número 1 (um) as entradas disponíveis, com 0 (zero) as não disponíveis e com as letras NA as entradas para as quais não se tem informação.

8.5. Utilização da ferramenta ColNucleo

Uma vez instaladas as ferramentas CAPFITOGEN e selecionada a ferramenta ColNucleo, será necessário definir uma série de parâmetros para que a programação R funcione corretamente.

8.5.1 Parâmetros iniciais definidos pelo usuário

8.5.1.1 Parâmetro: ruta

Explicação: Caminho é onde se encontram as ferramentas (documentos) CAPFITOGEN. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo: F:/CAPFITOGEN, C:/CAPFITOGEN, D:/MinhasFerramentas (MeusDocumentos)/CAPFITOGEN, etc.

8.5.1.2 Parâmetro: pasaporte

Explicação: Escrever o nome do arquivo que contém a tabela de passaporte em formato de texto, lembrando de incluir a extensão (.txt). Por exemplo, se o arquivo se chamar "tabela" deverá se escrever: "tabela.txt". Lembrar que este arquivo deverá ser guardado previamente na pasta "Passaporte", que faz parte do conjunto de pastas que compõem o diretório CAPFITOGEN. Esta tabela contém uma coluna adicional denominada "AVAILAB", que diz respeito à tabela de passaporte usada como modelo para outras ferramentas CAPFITOGEN. Esta coluna adicional indica a disponibilidade de cada entrada a ser selecionada para uma coleção núcleo/nuclear.

8.5.1.3 Parâmetro: geoqual

Explicação: Selecionar esta opção quando os dados de passaporte foram analisados pela ferramenta GEOQUAL e contém, portanto, 51 colunas (e não as 46 do modelo de tabela exclusivo para ColNucleo, sem haver sido submetida a GEOQUAL). Utilizar, portanto, a tabela resultado de GEOQUAL denominada PasaporteOriginalEvaluadoGEOQUAL.txt como tabela de passaporte no ponto (item) anterior.

8.5.1.4 Parâmetro: totalqual

Explicação: Se sua tabela de passaporte é proveniente de GEOQUAL e se deseja considerar uma qualidade mínima para que seus dados sejam incluídos

na análise, determinar o valor de TOTALQUAL100 a ser usado como limite. Serão admitidos valores entre 0 (qualidade nula) a 100 (máxima qualidade).

8.5.1.5 Parâmetro: mapaelc

Explicação: Indicar o nome do arquivo que contém o mapa ELC (produto da aplicação da ferramenta ELCmapas), o qual deverá encontrar-se na pasta ELCmapas, cuja pasta é uma das pastas que compõe o diretório CAPFITOGEN. Este mapa deverá estar no formato DIVA-GIS (extensão .grd, tal e qual como produzido por ELCmapas) e ser escrito com a extensão. Assim, se o nome do mapa for "mapa_elc_brasil", deverá se escrever "mapa_elc_brasil.grd".

8.5.1.6 Parâmetro: statelc

Explicação: Indicar o nome do arquivo que contém a tabela de estatísticas descritivas do mapa ELC produzido pela ferramenta ELCmapas (esta ferramenta costuma nomear este arquivo como "Estadist_ELC_" mais o nome do país ou região). Da mesma forma que no mapa ELC, este arquivo deverá estar situado na pasta ELCmapas. Também da mesma forma, o nome deverá ser escrito junto com a extensão. Neste caso, por ser uma tabela, a extensão será txt. Assim, se o arquivo se chamar "Estadist_ELC_brasil", deverá se escrever "Estadist_ELC_brasil.txt".

8.5.1.7 Parâmetro: distdup

Explicação: Determinar o valor da distância (em km) sob o qual se considerará que dois lugares de coleta, na realidade, representam uma mesma população (duplicado geográfico). O valor zero (valor mínimo e por omissão) dará menos prioridade na hora de selecionar as entradas com idênticas coordenadas. Na medida em que o valor da distância aqui estipulado aumente, um maior número de entradas se considerará como duplicados geográficos.

8.5.1.8 Parâmetro: porcol

Explicação: Corresponde a intensidade da amostragem. Indicar o tamanho desejável para a coleção núcleo/nuclear, expresso num percentual do

tamanho da coleção original (valores de 0 a 100). Por exemplo, se a coleção original contiver 2000 entradas e for desejável ter uma coleção núcleo/nuclear de 200 entradas, o valor a indicar aqui será 10, ou para uma coleção núcleo/nuclear de 300 entradas, o valor será 15.

8.5.1.9 Parâmetro: estratcol

Explicação: Selecionar a estratégia de fixar as cotas de representação de cada categoria ecogeográfica do mapa ELC. Poderá se escolher entre as estratégias "C" constante (a mesma cota para todas as categorias), "P" proporcional (cotas proporcionais à quantidade de entradas de cada categoria) ou "L" logarítmica (cotas proporcionais ao logaritmo da quantidade de entradas de cada categoria).

8.5.1.10 Parâmetro: availab

Explicação: Selecionar esta opção quando for desejável utilizar a coluna sobre disponibilidade de entradas a serem selecionadas com destino para uma coleção núcleo. Lembrar que para esta ferramenta, a tabela de passaporte inclui uma coluna denominada "AVAILAB", onde se indica a disponibilidade das entradas da coleção original para integrar coleções núcleo, através dos códigos 0 (entrada não disponível), 1 (entrada disponível) e NA (sem informação/não disponível). Caso não se deseje informação sobre disponibilidade, a seleção de entradas se realizará sobre o total das entradas. A disponibilidade é um critério próprio do curador e poderá ser determinado pelo número de sementes conservadas, sua germinação ou outros diversos fatores.

8.5.1.11 Parâmetro: resultados

Explicação: Introduzir o caminho da pasta onde deseja guardar os resultados das análises. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo C:/Resultados, D:/MinhasFerramentas(Meus Documentos)/Resultados, etc.

8.6. Resultados de ColNucleo

Uma vez finalizada a análise, ColNucleo produzirá um total de quatro tabelas, se o usuário indicou utilizar os dados de disponibilidade (parâmetro 8.5.1.10) ou três, no caso que tenha decidido não utilizar tal informação.

8.6.1.1 “CoreCollection.txt”. Esta tabela contém a tabela de passaporte das entradas que ColNucleo selecionou como coleção núcleo ecogeográfica, com uma coluna adicional na extrema direita da tabela denominada “BGcat”, que indica o grupo ou categoria do mapa ELC a que pertence a entrada, de acordo com a localização do local de coleta.

8.6.1.2 “CoreCollect_Properties.txt”. Esta tabela mostra vários dos parâmetros introduzidos pelo usuário e que ColNucleo utilizou para obter a coleção núcleo. Os campos que inclui são “Allocation_strategy” ou estratégia de afixação, “Sample_size” ou percentual de intensidade de amostragem, “Use_availability_data” ou se se usa ou não dados de disponibilidade, “No_access_sampled”, tamanho da coleção núcleo obtida e “No_access_to_be_multiplied” (só para o caso de usar dados de disponibilidade) ou número de entradas que não estão disponíveis ou não se tem informação sobre sua disponibilidade, e que deveriam estar para conseguir uma coleção núcleo completa.

8.6.1.3 “CoreCollect_stats.txt”. Esta tabela contém estatísticas por cada categoria do mapa ELC (identificadas na coluna “ELC_CAT”). Contém as seguintes colunas à direita de “ELC_CAT”: “FREC_W_DUPL” que indica o número de entradas, incluindo duplicados geográficos, cujo local de coleta ocorre dentro de cada categoria, “FREC_WO_DUPL” indica o mesmo que a anterior coluna mas excluindo duplicados geográficos, “Porcent_W_DUPL” indica o percentual de entradas (contando duplicados) que ocorre em cada categoria, “FreqClass_W_DUPL” indica a classificação em quartos da frequência de ocorrência em cada categoria de maneira similar a como se indica na seção 6.4.1.1, “Duplicates” é o número de entradas duplicados geográficos por cada categoria, “N_Availab” é o número de entradas disponíveis no total (duplicados mais não duplicados) para cada categoria, “N_AvailabWO” é o número de entradas disponíveis não duplicadas para cada categoria, “Q_Even” ou “Q_Prop” ou “Q_Log” (apareceria o título de coluna dependendo do método de afixação selecionado) é o número de

entradas que devem representar, a priori, cada categoria, ou seja, a cota, e, finalmente, a coluna CCfinal que corresponde ao número de entradas que compõem a coleção núcleo ecogeográfica obtida por ColNucleo com base nos parâmetros introduzidos e, neste caso, disponibilidade de entradas.

8.6.1.4 “EntriesToBeMultiplied.txt”. Se trata de uma tabela com a mesma estrutura de colunas que “CoreCollection.txt” só que em “EntriesToBeMultiplied.txt” aparecem as entradas que ColNucleo selecionou como parte da coleção núcleo, mas para as quais não há disponibilidade. Em princípio se assume, pelo nome do arquivo, que podem ser entradas que requeririam ser multiplicadas para ser disponíveis e, então, fazer parte da coleção núcleo. Contudo, a razão pelas quais não estão disponíveis podem ser variadas como se explica na seção 8.1.3.

8.7. Referências

- Brown, A.H.D. 1995. The core collection at the crossroads. p. 3–19. En Hodgkin, T., Brown, A.H.D., Hintum, T.J.L., Morales, E.A.V. (ed.) Core collections of plant genetic resources. John Wiley & Sons, New York, NY.
- Ghamkhar, K., R. Snowball, B.J. Wintle, Brown, A.H.D. 2008. Strategies for developing a core collection of bladder clover (*Trifolium spumosum* L.) using ecological and agro-morphological data. Aust. J. Agric. Res. 59:1103–1112.
- Grenier, C., Hamon, P., Bramel-Cox, P.J.. 2001. Core collection of sorghum: II. Comparison of three random sampling strategies. Crop Science. 41:241–246.
- Parra Quijano, M., Iriondo, J.M., Torres, M.E., De la Rosa, L. 2011a. Evaluation and validation of ecogeographical core collections using phenotypic data. Crop Science 51:694-703.
- Parra-Quijano, M., Iriondo, J.M., de la Cruz, M., Torres, M.E. 2011. Strategies for the development of core collections based on ecogeographical data. Crop Science 51:656-666.
- Reddy, L.J., H.D. Upadhyaya, C.L.L. Gowda, S. Singh. 2005. Development of core collection in pigeonpea (*Cajanus cajan* (L.) Millspaugh) using geographic and qualitative morphological descriptors. Genetic Resources and Crop Evolution 52:1049–1056.
- Tohme, J., P. Jones, S. Beebe, and M. Iwanaga. 1995. The combined use of agroecological and characterisation data to establish the CIAT Phaseolus vulgaris core collection. p. 95–107. In Hodgkin, T., Brown, A.H.D., Hintum, T.J.L., Morales, E.A.V. (ed.) Core collections of plant genetic resources. John Wiley & Sons, New York, NY.
- Upadhyaya, H.D., Ortiz, R., Bramel, P.J., S. Singh, S. 2003. Development of a groundnut core collection using taxonomical, geographical and morphological descriptors. Genet. Resour. Crop Evol. 50:139–148.
- Xiurong, Z., Yingzhong, Z., Yong, C., Xiangyun, F., Qingyuan, G., Mingde, Z., Hodgkin, T. 2000. Establishment of sesame germplasm core collection in China. Genet. Resour. Crop Evol. 47:273–279.

Yonezawa, K., Nomura, T., Morishima, H. 1995. Sampling strategies for use in stratified germplasm collections. p. 35–53. In Hodgkin, T., Brown, A.H.D., Hintum, T.J.L., Morales, E.A.V. (ed.) Core collections of plant genetic resources. John Wiley & Sons, New York, NY.

9. Ferramenta FIGS R

9.1. A Estratégia de Seleção Focada de Germoplasma

A técnica de seleção de germoplasma com fins de utilização denominada “Estratégia de Seleção Focada de Germoplasma” ou FIGS por suas siglas em inglês (Focused Identification of Germplasm Strategy), provém de um conceito originalmente desenvolvido por Mackay (1990).

Esta técnica busca identificar entradas de uma coleção que têm potencial de uso por parte de melhoradores. O potencial para uso em melhoramento está baseado no uso de informação ecogeográfica dos locais de coleta e sua associação com caracteres de interesse para o melhorador de cultivos (Mackay e Street, 2004).

Dado que FIGS usa variáveis ecogeográficas de tipo abiótico para a seleção de germoplasma, a associação entre as variáveis ecogeográficas e o carácter de interesse para o melhoramento é direta, se o carácter de interesse é de tipo abiótico ou, indireto, para um carácter biótico. Desta maneira, se um melhorador busca germoplasma com potencial e o carácter de interesse é adaptação a condições de seca, buscará diretamente germoplasma cujo local de coleta esteja caracterizado por baixas precipitações. Se a característica de interesse fosse de tipo biótica, como a resistência a um patógeno, primeiro se deveria estabelecer uma relação entre uma série de variáveis ecogeográficas e a resistência a tal patógeno, para, conseqüentemente, selecionar germoplasma cujo local de coleta reúna as condições ecogeográficas associadas à resistência ao patógeno.

Se podem distinguir duas técnicas para a seleção de germoplasma mediante FIGS. A primeira é a de filtragem de entradas e a segunda é a técnica de calibragem.

A técnica de filtragem seleciona entradas de uma coleção caracterizada ecogeograficamente, escolhendo aquelas que cumprem com certos valores ou categorias para alguma das variáveis pelas quais foram caracterizadas. Em algumas ocasiões, o que se seleciona é uma fração da distribuição de uma variável ecogeográfica na coleção caracterizada. Os valores, categorias ou a fração da distribuição e a variável ecogeográfica de seleção são determinados

pelo investigador, curador ou melhorador, com base no conhecimento da espécie, da variável ecogeográfica e do carácter de interesse. Um exemplo da aplicação deste método é a seleção FIGS indireta para resistência do trigo frente à praga *Eurygaster integriceps* (El Bouhssini et al., 2009). Outro caso é a aplicação de FIGS de forma direta para a identificação de recursos genéticos de *Vicia faba* com adaptação à seca (Khazaei et al., 2013).

A técnica de calibragem requer que a coleção tenha sido caracterizada ecogeograficamente em sua totalidade ou quase totalidade (entradas com coordenadas) e que, além disso, se tenha avaliado ao menos parcialmente pelo carácter de interesse. A técnica de calibragem se desenvolve em duas fases. Na primeira fase, se utiliza análises matemáticas e estatísticas para estabelecer a relação entre a presença ou ausência do carácter de interesse e uma ou várias variáveis ecogeográficas. Uma vez provada esta relação, se realiza um prognóstico de presença ou ausência do carácter de interesse sobre a fração da coleção não avaliada usando, para isso, a informação ecogeográfica disponível para a totalidade da coleção. O prognóstico assinalaria quais entradas seriam potencialmente de interesse para o melhoramento de cultivos. A aplicação da técnica de calibragem se observa nos estudos de Endresen e colaboradores para cevada e trigo (Endresen, 2010; Endresen et al., 2012).

Por sua própria natureza, a técnica de calibragem se usa para FIGS de tipo indireto, enquanto a técnica de filtragem se tem usado para ambos tipos. A técnica de calibragem é metodologicamente mais complexa, e seus resultados são mais precisos na detecção de entradas que possuem a característica de interesse com respeito à técnica de filtragem. Contudo, a técnica de calibragem tem como desvantagem, o requisito de ter que contar com dados de avaliação parcial da coleção e que esses dados sejam suficientemente fiáveis, como para que a relação que se possa estabelecer entre variável ecogeográfica e carácter de interesse seja válida. Este fato faz com que sua aplicação fique restrita a 22% das coleções, que é o percentual de coleções nacionais que contam com algum tipo de avaliação por fatores bióticos em 40 países, de acordo com o Segundo Relatório do Estado dos Recursos Fitogenéticos para a Alimentação e a Agricultura (FAO, 2010).

Independentemente da forma como se obtenha a subcoleção FIGS, é altamente recomendável que se valide mediante ensaios de adaptação, tolerância ou resistência, que efetivamente as entradas selecionadas possuam a característica de interesse pela qual foram selecionadas, através das condições ecogeográficas de seus locais de coleta.

9.2. Subcoleções FIGS e coleções núcleo

Uma subcoleção FIGS é o conjunto de entradas com potencial uso no melhoramento de uma espécie cultivada e que provém de uma seleção FIGS.

As subcoleções FIGS, à diferença de uma coleção núcleo, não têm necessariamente que ser representativas da variabilidade da coleção original. Uma subcoleção FIGS convencional implica uma tendência ou desvio ou vício (ou interesse do melhorista de plantas) muito particular no momento de sua seleção, motivo pelo qual não cabe esperar encontrar alta representatividade.

Outra diferença entre uma coleção núcleo e uma subcoleção FIGS é que para esta última, pode-se estabelecer por espécie, tantas quantas características de interesse existam. Em contraste, coleções núcleo geralmente são estabelecidas uma por espécie.

Contudo, como no caso das coleções núcleo, estabelecer uma ou várias subcoleções FIGS não deve pôr em risco a conservação das entradas não selecionadas. Se uma coleção núcleo serve, por exemplo, para priorizar alguns trabalhos como a caracterização ou avaliação de entradas numa coleção quando não se conta com meios para realizar o trabalho para a totalidade da coleção, uma subcoleção FIGS o que busca é potencializar a utilização do germoplasma contido numa coleção, facilitando ao melhorista de plantas, que é seu principal usuário, a busca de material com potencial para ser integrado em seus programas de melhoramento.

9.3. Obtenção de subcoleções FIGS com a ferramenta FIGS_R

Com a ferramenta FIGS_R se pode obter uma subcoleção FIGS mediante a técnica de filtragem. FIGS_R permite utilizar até três variáveis de seleção de maneira hierárquica, isto é, uma variável primária (de uso obrigatório) que exerce o primeiro processo de filtragem, uma secundária (opcional) que filtra sobre o

subconjunto produto da primeira filtragem e uma variável terciária (opcional e só funcional se se utilizou uma variável secundária) que filtra sobre o subconjunto produto da segunda filtragem. Como variável primária, secundária ou terciária, se pode escolher uma das 103 variáveis ecogeográficas (ver Anexos 12.1, 12.2 e 12.3) que se têm à disposição nas ferramentas CAPFITOGEN. A Fig. 31 mostra o processo que segue FIGS_R para a obtenção de subcoleções FIGS.

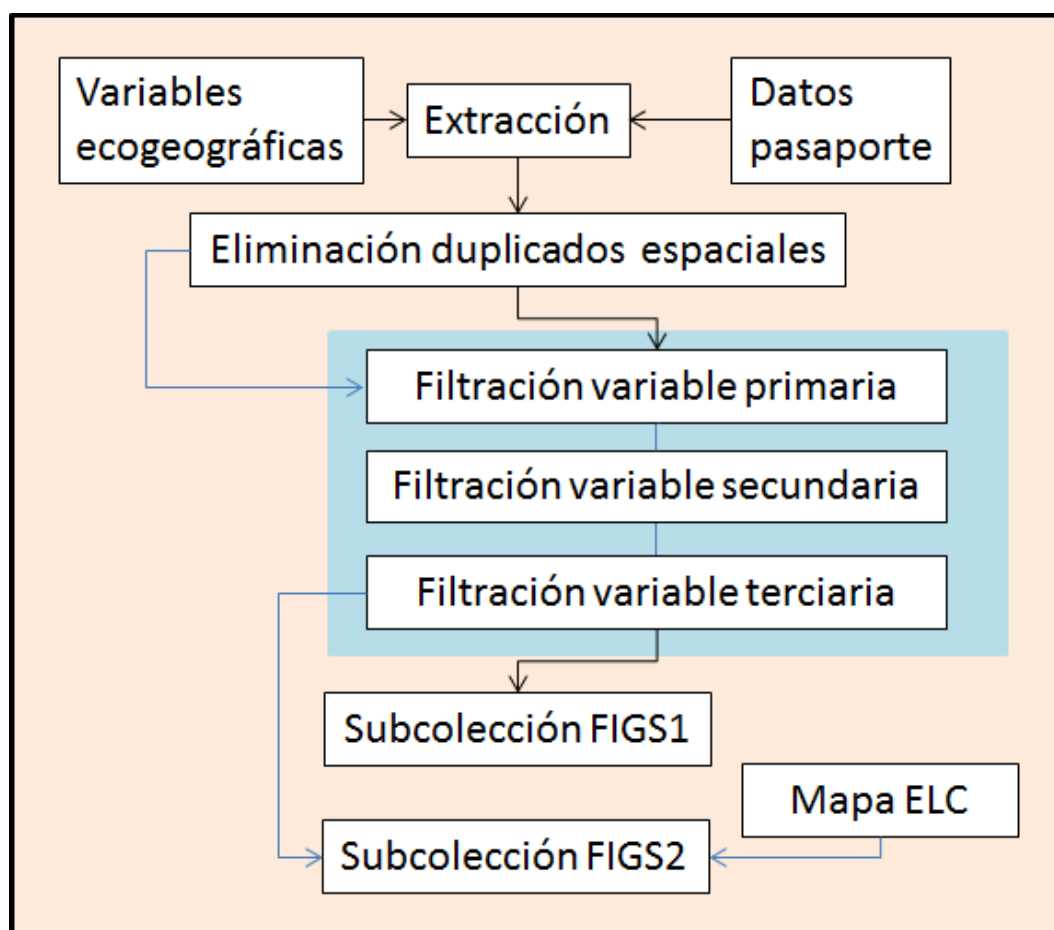


Figura 31. Ilustração do processo que segue a ferramenta FIGS_R para a obtenção de subcoleções FIGS.

Para cada variável de seleção, o usuário da ferramenta FIGS_R pode escolher a forma de realizar a seleção. A primeira forma é a da determinação por parte do usuário de uma categoria de valores que deve satisfazer uma entrada para ser incluída na subcoleção FIGS. A segunda forma é a determinação de uma fração da coleção, expressa em percentual, cujas entradas são as que possuem os valores mais altos ou mais baixos em relação à variável de seleção.

FIGS_R adota alguns termos ou definições do melhoramento genético de plantas, como intensidade de seleção ou diferencial de seleção. O primeiro se usa para definir o percentual da coleção inicial que ficará incluído na subcoleção FIGS. O segundo, faz referência à diferença entre a média da coleção original e a subcoleção FIGS para a ou as variáveis de seleção.

Além disso, a ferramenta FIGS_R permite criar subcoleções FIGS balanceadas ecogeograficamente. Em outras palavras, se o usuário previamente gerou um mapa ELC (através da ferramenta ELC mapas, capítulo 4) e usa exclusivamente a segunda forma de seleção (fração da coleção), pode fazer que a ferramenta: 1. Atribua categorias a cada entrada com base na categoria do mapa ELC onde ocorre seu local de coleta, e 2. selecione de cada categoria a fração de entradas com os valores mais altos ou baixos para a variável de seleção. O balanço mediante um mapa ELC tende a produzir subcoleções FIGS com uma maior representatividade ecogeográfica, ainda que conservando em boa medida o potencial de utilização em programas de melhoramento pela característica de interesse.

Finalmente, é importante considerar que a ferramenta FIGS_R pode levar em conta, também, informação sobre disponibilidade das entradas a serem selecionadas. Desta maneira, utiliza o mesmo formato de entrada de dados (dados de passaporte) que a ferramenta ColNucleo, ou seja, o formato usado para a ferramenta GEOQUAL mais a adição do campo “AVAILAB”. Para preencher o campo “AVAILAB” se aplicam os mesmos critérios descritos na seção 8.4.

9.4. Utilização da ferramenta FIGS_R

Uma vez instaladas as ferramentas CAPFITOGEN e selecionada a ferramenta FIGS_R, será necessário definir uma série de parâmetros para que a programação R funcione corretamente.

9.4.1 Parâmetros iniciais definidos pelo usuário

9.4.1.1 Parâmetro: ruta

Explicação: Caminho é onde se encontram as ferramentas (documentos) CAPFITOGEN. Nota: usar / em vez de \ na indicação do caminho da pasta. Por

exemplo: F:/CAPFITOGEN, C:/CAPFITOGEN, D:/MinhasFerramentas (MeusDocumentos)/CAPFITOGEN, etc.

9.4.1.2 Parâmetro: pais

Explicação: Selecionar o país no qual foram recoletadas a maioria ou a totalidade das entradas que se deseja analisar. Se as entradas foram recoletadas em mais de um país, considerar selecionar uma região, subcontinente ou continente (progressivamente serão adicionadas estas opções).

9.4.1.3 Parâmetro: pasaporte

Explicação: Escrever o nome do arquivo que contém a tabela de passaporte em formato de texto, lembrando de incluir a extensão (.txt). Por exemplo, se o arquivo se chamar "tabela" deverá se escrever: "tabela.txt". Lembrar que este arquivo deverá ser guardado, previamente, na pasta "Passaporte", que formará parte do conjunto de pastas que compõem o diretório CAPFITOGEN. Esta tabela contém uma coluna adicional denominada "AVAILAB", relacionada à tabela de passaporte usada como modelo para outras ferramentas CAPFITOGEN. Esta coluna adicional indica a disponibilidade de cada entrada a ser selecionada para uma subcoleção FIGS.

9.4.1.4 Parâmetro:geoqual

Explicação: Selecionar esta opção quando os dados de passaporte foram analisados pela ferramenta GEOQUAL e contém, portanto, 51 colunas (e não as 46 do modelo de tabela exclusivo para ColNucleo, sem haver sido submetida a GEOQUAL). Sendo assim, usar a tabela resultado de GEOQUAL, denominada PasaporteOriginalEvaluadoGEOQUAL.txt, como tabela de passaporte no ponto (item) anterior.

9.4.1.5 Parâmetro: totalqual

Explicação: Se sua tabela de passaporte for proveniente de GEOQUAL e for desejável considerar um mínimo de qualidade para que seus dados possam ser incluídos na análise, determinar o valor de TOTALQUAL100 a ser usado

como limite. Admitir-se-á valores de 0 (qualidade nula) a 100 (máxima qualidade).

9.4.1.6 Parâmetro: controlelc

Explicação: Indicar se for desejável usar um mapa ELC criado previamente para determinar a distribuição das entradas que compõem a subcoleção FIGS, sobre as diferentes categorias do mapa. Para usuários avançados (com maior experiência), esta opção habilitará a obtenção de uma subcoleção FIGS adicional, onde a seleção das entradas se fará sobre cada categoria ELC. Para isto será necessário usar métodos de seleção sobre frações de distribuição para todas as variáveis consideradas.

9.4.1.7 Parâmetro: mapaelc

Explicação: Indicar o nome do arquivo que contém o mapa ELC (produto da aplicação da ferramenta ELCmapas), o qual deverá encontrar-se na pasta ELCmapas, que por sua vez é uma das pastas que compõem o diretório CAPFITOGEN. Este mapa deverá estar no formato DIVA-GIS (extensão .grd tal como produzido pelo ELCmapas) e deverá ser escrito com a extensão. Assim, se o nome do mapa é "mapa_elc_brasil", deverá ser escrito "mapa_elc_brasil.grd".

9.4.1.8 Parâmetro: statelc

Explicação: Indicar o nome do arquivo que contém a tabela de estatísticas descritivas do mapa ELC, produzido pela ferramenta ELCmapas (esta ferramenta costuma nomear este arquivo como "Estadist_ELC_" mais o nome do país ou região). Da mesma forma que para o mapa ELC, este arquivo deverá estar situado na pasta ELCmapas. O nome deverá ser escrito junto com a extensão e, neste caso, por tratar-se de uma tabela, a extensão será txt. Assim, se o arquivo se chamar "Estadist_ELC_brasil" deverá ser escrito "Estadist_ELC_brasil.txt".

9.4.1.9 Parâmetro: distdup

Explicação: Determinar o valor da distância (em km) sob a qual se considerará que dois lugares de recoleção, na realidade, representam uma mesma

população (duplicado geográfico). O valor zero (valor mínimo e por omissão) dará menor prioridade na hora de selecionar as entradas com idênticas coordenadas. Na medida em que o valor da distância aqui estipulado aumentar, um maior número de entradas será considerado como duplicado geográfico.

9.4.1.10 Parâmetro: availab

Explicação: Selecionar esta opção quando for desejável utilizar a coluna sobre disponibilidade de entradas a serem selecionadas como integrantes da subcoleção FIGS. Isto implica em dar prioridade à entradas disponíveis, sem restringir uma possível consideração das não disponíveis. Lembrar que para esta ferramenta, a tabela de passaporte inclui uma coluna denominada "AVAILAB", onde se indica a disponibilidade das entradas da coleção original para integrar as coleções núcleo, através dos códigos 0 (entrada não disponível), 1 (entrada disponível) e NA (sem informação/não disponível). A disponibilidade é um critério próprio do curador e poderá ser determinado pelo número de sementes conservadas, sua germinação ou diversos outros fatores.

9.4.1.11 Parâmetro: soloavailab

Explicação: Selecionar esta opção quando for desejável restringir a seleção de entradas com destino para subcoleção FIGS, exclusivamente para as entradas designadas como disponíveis (valor 1 no campo "AVAILAB").

9.4.1.12 Parâmetro: resol1

Explicação: Selecionar o nível de resolução que se deseja utilizar para a extração da informação ecogeográfica. Note que 1x1 km oferece maior resolução, mas exige maior capacidade computacional em relação a 5x5 km, ainda que este não seja um aspecto tão limitante como na ferramenta ELC mapas. As resoluções 10x10 e 20x20 km se restringem a países de grande extensão territorial, subcontinentes ou continentes.

9.4.1.13 Parâmetro:buffy

Explicação: Marcar esta opção quando for desejável que a extração da informação ecogeográfica seja feita sobre uma área ao redor do lugar de

recolecção. Deixar de marcar esta opção fará com que a extração seja realizada somente para o ponto que indica as coordenadas de lugar de recolecção.

9.4.1.14 Parâmetro: tamp

Explicação: Especificar o raio (em metros) de uma área circular ao redor do ponto que indica as coordenadas do lugar de recolecção, do qual se extrairá a informação ecogeográfica. Para se obter um único valor, se usarão médias dos valores extraídos da área circular.

9.4.1.15 Parâmetro: variab1v

Explicação: Selecionar uma (1) variável ecogeográfica primária para qual se deseja selecionar entradas com a finalidade de se obter uma subcoleção FIGS. Se for preferível selecionar entradas para uma ou duas variáveis adicionais (variável secundária e terciária), a variável selecionada neste ponto será aquela com a qual se realizará o primeiro filtrado.

9.4.1.16 Parâmetro: variab1rang

Explicação: Marcar esta opção quando for desejável selecionar entradas que digam respeito à variável primária usando um intervalo de valores, ou seja, indicando um valor máximo e mínimo que determine uma categoria que será usada para selecionar as entradas com destino à subcoleção FIGS.

9.4.1.17 Parâmetro: variab1min

Explicação: Especificar o valor mínimo para a variável primária com a qual se determinará a categoria que será usada para selecionar entradas com destino à subcoleção FIGS.

9.4.1.18 Parâmetro: variab1max

Explicação: Especificar o valor máximo para a variável primária com a qual será determinada a categoria que se usará para selecionar entradas com destino à subcoleção FIGS.

9.4.1.19 Parâmetro: variab1cola

Explicação: Marcar esta opção se for desejável selecionar entradas que digam respeito à variável primária usando uma fração da distribuição, ou seja, um percentual da coleção original cujos valores serão os mais altos ou os mais baixos em relação à variável primária.

9.4.1.20 Parâmetro: variab1vpor

Explicação: Determinar a fração da distribuição (em percentual) que se deseja selecionar para formar a subcoleção FIGS. Valores permitidos entre 0 e 100.

9.4.1.21 Parâmetro: variab1vhl

Explicação: Selecionar a fração da distribuição que se deseja selecionar com referência à variável primária.

9.4.1.22 Parâmetro: variab2

Explicação: Marcar esta opção se for desejável utilizar uma variável secundária para selecionar entradas com destino à uma subcoleção FIGS. Com os valores desta variável será realizada uma seleção de entradas do subconjunto previamente selecionado com a variável primária.

9.4.1.23 Parâmetro: variab2v

Explicação: Selecionar uma (1) variável ecogeográfica secundária para a qual se deseja selecionar entradas com a finalidade de obter uma subcoleção FIGS. Poderá ser a mesma variável primária.

9.4.1.24 Parâmetro: variab2rang

Explicação: Marcar esta opção quando for desejável selecionar entradas que digam respeito à variável secundária usando um intervalo de valores, ou seja, indicando um valor máximo e mínimo que determine uma categoria que será usada para selecionar as entradas com destino à subcoleção FIGS.

9.4.1.25 Parâmetro: variab2min

Explicação: Especificar o valor mínimo para a variável secundária com a qual se determinará a categoria que será usada para selecionar entradas com destino à subcoleção FIGS.

9.4.1.26 Parâmetro: variab2max

Explicação: Especificar o valor máximo para a variável secundária com a qual se determinará a categoria que será usada para selecionar entradas com destino à subcoleção FIGS.

9.4.1.27 Parâmetro: variab2cola

Explicação: Marcar esta opção quando for desejável selecionar entradas que digam respeito à variável secundária usando uma fração da distribuição restante, ou seja, um percentual do subconjunto selecionado para a variável primária, cujos valores serão os mais altos ou os mais baixos em relação a variável secundária.

9.4.1.28 Parâmetro: variab2vpor

Explicação: Determinar a fração da distribuição restante (em percentual) que se deseja selecionar para formar a subcoleção FIGS, usando a variável secundária. Valores permitidos entre 0 e 100.

9.4.1.29 Parâmetro: variab2vhl

Explicação: Selecionar a fração da distribuição que se deseja com relação à variável secundária.

9.4.1.30 Parâmetro: variab3

Explicação: Marcar esta opção quando for desejável utilizar uma variável terciária para selecionar entradas com destino à uma subcoleção FIGS. Com os valores desta variável será realizada uma seleção de entradas sobre o subconjunto previamente selecionado para as variáveis primária e secundária. Caso não seja determinada previamente a utilização de uma variável secundária, selecionar uma terciária não trará nenhum efeito para a formação da subcoleção FIGS.

9.4.1.31 Parâmetro: variab3v

Explicação: Selecionar uma (1) variável ecogeográfica terciária para a qual se deseja selecionar entradas com a finalidade de obter uma subcoleção FIGS. Poderá ser a mesma variável primária ou secundária.

9.4.1.32 Parâmetro: variab3rang

Explicação: Marcar esta opção quando for desejável selecionar entradas que digam respeito à variável terciária usando um intervalo de valores, ou seja, indicando um valor máximo e mínimo que determine uma categoria que será usada para selecionar as entradas com destino à subcoleção FIGS.

9.4.1.33 Parâmetro: variab3min

Explicação: Especificar o valor mínimo para a variável terciária, com a qual se determinará a categoria que será usada para selecionar entradas com destino à subcoleção FIGS.

9.4.1.34 Parâmetro: variab3max

Explicação: Especificar o valor máximo para a variável terciária, com a qual se determinará a categoria que será usada para selecionar entradas com destino à subcoleção FIGS.

9.4.1.35 Parâmetro: variab3cola

Explicação: Marcar esta opção quando for desejável selecionar entradas que digam respeito à variável terciária usando uma fração da distribuição restante, ou seja, um percentual do subconjunto selecionado para as variáveis primária e secundária, cujos valores serão os mais altos ou os mais baixos em relação à variável terciária.

9.4.1.36 Parâmetro: variab3vpor

Explicação: Determinar a fração da distribuição restante (em percentual) que se deseja selecionar para formar a subcoleção FIGS, usando a variável terciária. Valores permitidos entre 0 e 100.

9.4.1.37 Parâmetro: variab3vhl

Explicação: Selecionar a fração da distribuição que se deseja selecionar em relação à variável terciária.

9.4.1.38 Parâmetro: resultados

Explicação: Introduzir o caminho da pasta onde deseja guardar os resultados das análises. Nota: usar / em vez de \ na indicação do caminho da pasta. Por exemplo C:/Resultados, D:/MinhasFerramentas(Meus Documentos)/Resultados, etc.

9.5. Resultados de FIGS_R

Uma vez finalizada a análise e de acordo com a configuração dos parâmetros anteriormente detalhados, FIGS_R produzirá entre três e cinco tabelas.

Quando não se inclui um mapa ELC na análise (parâmetro 9.4.1.6, controlelc), só aparecerão as seguintes três tabelas:

9.5.1 “FIGS_regular.txt”. Esta tabela contém a identificação das entradas selecionadas com destino à subcoleção FIGS (campo “ACCENUMB”), as coordenadas do local de coleta (“DECLATITUDE” e “DECLONGITUDE”), o campo de disponibilidade (“AVAILAB”) e, à continuação, tantas colunas como variáveis de seleção usadas.

9.5.2 “FIGS_stat_table.txt”. Corresponde à tabela que resume as características tanto da coleção original como da subcoleção FIGS através de estatísticas, como a intensidade de seleção alcançada, média, valor máximo e mínimo e diferencial de seleção para cada uma das variáveis de seleção utilizadas.

9.5.3 “Passport_FIGS_R.txt”. Esta tabela corresponde à tabela de passaporte em que o usuário introduz na análise um campo adicional para cada variável de seleção utilizada, chamado “SEL_VAR”, seguido pelo número 1, 2 ou 3. Estes campos assinalam com o número 1 as entradas que fazem parte da subcoleção FIGS e, com “NA”, as que não foram selecionadas. Assim, no processo seletivo com a variável primária (definida no parâmetro 9.4.1.15) se teria selecionado as entradas identificadas com “1” no campo “SEL_VAR1”, se foi utilizada uma variável de seleção secundária; as entradas selecionadas no

segundo processo de filtragem aparecerão identificadas com “1” no campo “SEL_VAR2”, e, finalmente, se foi utilizada uma variável de seleção terciária, as entradas selecionadas no terceiro processo de filtragem aparecerão identificadas com “1” no campo “SEL_VAR3”.

Quando se inclui um mapa ELC para conhecer as características ecogeográficas da subcoleção FIGS gerada, aparecerá uma nova tabela:

9.5.4 “FIGS_freq_ELCmap.txt”. Esta tabela mostra valores de frequência, número de duplicados e número de entradas disponíveis para cada categoria ecogeográfica, de maneira similar à tabela da ferramenta ColNucleo, descrita na seção 8.6.1.3 (“CoreCollect_stats.txt”). Além disso, aparecem na parte esquerda da tabela até três novos campos identificados com o prefixo “FIGS_var” e, à continuação, os números 1, 2 ou 3. Desta maneira, no campo “FIGS_var1” aparecerá o número de entradas selecionadas pela variável primária que foram coletadas em cada categoria ELC; em “FIGS_var2”, o número de entradas selecionadas pela variável secundária no segundo processo de filtragem de cada categoria ELC e, em “FIGS_var3”, o número de entradas selecionadas pela variável terciária no terceiro processo de filtragem de cada categoria ELC.

Finalmente, se para as variáveis de seleção primária, secundária e terciária se utilizou exclusivamente a segunda forma de seleção (fração da coleção), isto é, se foi marcada a opção “variab1cola”, “variab2cola” e “variab3cola” (parâmetros 9.4.1.19, 9.4.1.27 e 9.4.1.35, respectivamente), os resultados incluirão uma quinta tabela:

9.5.5 “FIGS_UnderELC.txt”. Corresponde a uma tabela com os mesmos campos descritos anteriormente para a tabela “Passport_FIGS_R.txt” (seção 9.5.3) mas, neste caso, só contém as entradas que constituem a subcoleção FIGS balanceada mediante o mapa ELC, e incluem os campos “SEL_VAR1”, “SEL_VAR2” e “SEL_VAR3”, que indicam com o valor 1 se essas entradas foram também selecionadas num esquema FIGS sem o uso de mapas ELC. Além disso, no lado esquerdo da tabela aparecem até 3 novos campos, identificados como “var_eco1”, “var_eco2” e “var_eco3”, dependendo de quantas variáveis de seleção se tenham utilizado. Em cada um destes campos

aparecerão os valores correspondentes à extração de cada variável de seleção para cada local de coleta (“var_eco1” valores da variável primária; “var_eco2” valores da variável secundária; e “var_eco3” valores da variável terciária).

Além disso, a tabela “FIGS_freq_ELCmap.txt” (seção 9.5.4) incluirá até três novos campos no lado esquerdo, com as denominações “No_by_var1”, “No_by_var2” e “No_by_var3”. Nestes campos se mostra o número de entradas selecionadas para a subcoleção FIGS balanceada pelo mapa ELC em cada processo de seleção. “No_by_var1” para o primeiro processo de filtragem pela variável primária; “No_by_var2” para o segundo processo de filtragem pela variável secundária; e “No_by_var3” para o terceiro processo de filtragem pela variável terciária.

9.6. Referências

- El Bouhssini, M. E., Street, K., Joubi, A., Ibrahim, Z., Rihawi, F. 2009. Sources of wheat resistance to Sunn pest, *Eurygaster integriceps* Puton, in Syria. Genetic Resources and Crop Evolution 56: 1065–1069.
- Endresen, D.T.F. 2010. Predictive association between trait data and ecogeographic data for Nordic barley landraces. Crop Science 50: 2418-2430.
- Endresen, D.T.F., Street, K., Mackay, M., Bari, A., Amri, A., De Pauw, E., Nazari, K., Yahyaoui, A. 2012. Sources of resistance to stem rust (Ug99) in bread wheat and durum wheat identified using Focused Identification of Germplasm Strategy. Crop Science 52: 764-773.
- FAO 2010. The Second Report on the State of the World's Plant Genetic Resources for Food and Agriculture. Rome
- Khazaei, H., Street, K., Bari, A., Mackay, M., Stoddard, F.L. 2013. The FIGS (Focused Identification of Germplasm Strategy) approach identifies traits related to drought adaptation in *Vicia faba* genetic resources. PLoS ONE 8(5): e63107. doi:10.1371/journal.pone.0063107
- Mackay, M.C. 1990. Strategic planning for effective evaluation of plant germplasm. p. 21-25 En: Srivastava, J.P., Damania, A.B. (eds). Wheat genetic resources: Meeting diverse needs. John Wiley & Sons, Chichester, UK.
- Mackay, M. C., Street, K. 2004. Focused identification of germplasm strategy – FIGS. p. 138-141. En: Black, C.K., Panozzo, J.F., Rebetzke, G.J. (eds). Cereals 2004. Proceedings of the 54th Australian Cereal Chemistry Conference and the 11th Wheat Breeders' Assembly, 21-24 September 2004, Canberra, Australian Capital Territory (ACT). Cereal Chemistry Division, Royal Australian Chemical Institute, Melbourne, Australia.

10. Erros freqüentes

Na seguinte lista se encontram muitos dos avisos de erro (texto de cor vermelha na parte inferior da interface), ou ao menos os mais freqüentes, que podem aparecer durante a execução de alguma das ferramentas. Se durante sua experiência de trabalho com as ferramentas CAPFITOGEN aparecem avisos de erro diferentes aos expostos ou se se tem alguma dúvida sobre o funcionamento das mesmas, pode-se visitar o foro de resolução de problemas no seguinte link: <http://www.agrobiodiversidad.org/foroCAPFITOGEN/>.

1. Sem aviso de erro mas sem resultados na pasta designada para isso:

Ferramenta: Qualquer pois provém de uma localização inadequada do conjunto de ferramentas

Solução(ões): Verifique se a estrutura de dados das ferramentas (pastas e arquivos incluídos na pasta chamada CAPFITOGEN) se encontram dentro de um caminho sem valores atípicos ou sem espaços. Por exemplo, se você situou as ferramentas na pasta “Mis documentos”, o caminho poderia ser assim: C:\Mis documentos\CAPFITOGEN. Este caminho contém um espaço entre as palavras “Mis” e “documentos” o qual pode gerar este erro, onde aparentemente a análise se executou corretamente mas a pasta designada para guardar os resultados está vazia. Recomenda-se localizar a pasta CAPFITOGEN diretamente na raiz do diretório da unidade de disco (neste caso assim: C:\CAPFITOGEN).

2. Aviso de erro: An error occurred: Error en library(package name) : there is no package called 'package name'Calls: source -> withVisible -> eval -> eval -> library

Ferramenta: Qualquer pois provém da instalação

Solução(ões): Este erro indica que algum pacote de R que requer a ferramenta não foi corretamente instalado e por esa razão R não o encontra. Assegure-se de que a estrutura de pastas e arquivos das ferramentas não se encontra localizada na raiz do diretório (por exemplo em K:/). Se é assim, crie uma pasta (normalmente com o nome de CAPFITOGEN) na raiz do diretório, copie toda a estrutura de pasta e arquivos e cole-a na pasta que acaba de criar. À continuação, reinstale as ferramentas. Se esta opção não funciona, prove instalar o pacote manualmente. Para isso:

a. No código do erro aparece o nome do pacote não instalado, exatamente onde no exemplo diz ‘package name’. Com este nome de pacote, abra a pasta “packages” dentro da estrutura de pastas e arquivos de CAPFITOGEN e aí se encontrará uma série de arquivos .zip com diversos nomes. Um deles corresponderá ao nome do pacote e estará acompanhado dos números relativos à versão. Copie o nome do arquivo (completo, incluindo a extensão .zip).

b. Abra R, com o executável que estará alojado no caminho: C:\rwin\bin\i386\Rgui.exe e digite a seguinte instrução na “R console”:

Install.packages(“X:/CAPFITOGEN/packages/nombreadchivopaquete.zip”)

X corresponde à letra da unidade onde se encontram as ferramentas CAPFITOGEN (troque esta letra com a que esteja de acordo com seu caso) e na parte “nombreadchivo.zip” cole o nome do arquivo que copiou no passo 1. À continuação digite “enter”.

c. O programa começará a instalar o pacote e quando acabar aparecerá um aviso similar a este: package ‘sp’ successfully unpacked and MD5 sums checked

d. Prove que o pacote ficou corretamente instalado, para isso digite à continuação:

library(“package name”)

“package name” é o nome do pacote e como aparece no aviso de erro (sem códigos de versões nem extensión .zip). Posteriormente pulse “enter”. Aparecerá um aviso similar a este:

Mensagens de aviso perdidos

package ‘cluster’ was built under R version 2.15.3

e. Prove de novo a ferramenta. É possível que apareça o mesmo erro mas para um pacote diferente. Se assim for, repita a operação até que cessem os avisos de erro.

Estes tipos de erros não são comuns desde que se melhorou o sistema de instalação, mas ocasionalmente se apresentam, especialmente na versão de Windows 8.

3. Aviso de erro: An error occurred: Error: no se puede ubicar un vector de tamaño X.X Gb

Ferramenta(s): várias

Solução(ões): É um problema relacionado com o tamanho das matrizes que tem que manejar R. Costuma-se solucionar rebaixando resoluções dos mapas. Se se apresenta em ELC mapas, se recomenda mudar o método de determinação de número ótimo de grupos ou incrementar o tamanho de célula no parâmetro resol1.

Também é possível que este erro apareça em GEOQUAL por um erro no conteúdo das tabelas de passaporte, especificamente quando ocorrem duplicados para o campo ACCENUMB. Este campo identifica inequivocamente cada entrada e um só duplicado pode gerar este aviso de erro. A solução está em revisar na tabela que não ocorram duplicados e se ocorrem, atribuir números ou códigos únicos para cada entrada.

4. Aviso de erro: An error occurred: Error en sample.int(m, k) : primer argumento inválidoCalls: source ... withVisible -> eval -> eval -> kmeans -> sample.int

Ferramenta(s): ELC mapas

Solução(ões): Significa que alguma variável para essa região ou país é constante e ao se estandarizar produz uma tabela de 0 filas que logo gera o erro em Kmeans (método elbow). Soluciona-se desseleccionando essa variável problemática. Para identificá-la, é importante saber que geralmente corresponde a variáveis de solo, especialmente em países pequenos. Por exemplo, a variável “profundidade” costuma originar este problema. Também em países secos, usar variáveis de mínimas precipitações costuma produzir este problema

5. Aviso de erro: An error occurred: Error en clara(sdata, k, ...) : x is not a numeric dataframe or matrix.Calls: source -> withVisible -> eval -> eval -> pamk -> clara

Ferramenta(s): ELC mapas

Solução(ões): Significa que alguma variável para essa zona é constante e ao se estandarizar produz uma tabela de 0 filas que logo gera o erro em medoides. Igual solução que o erro No. 2.

6. Aviso de erro: An error occurred: Error en kmeans(edaph[, -1], centers = i) : more cluster centers than distinct data points.Calls: source -> withVisible -> eval -> eval -> kmeans

Ferramenta(s): ELC mapas

Solução(ões): Significa que se indicou um número máximo de grupos menor ao número objetivo que o método elbow determina como ótimo. Repete-se a operação com um número menor de grupos.

7. Aviso de erro: An error occurred: Error: objeto 'ecogeot' no encontrado

Ferramenta(s): ECOGEO

Solução(ões): Deve-se seleccionar a opção geophysv se se seleccionam variáveis geofísicas.

8. Aviso de erro: An error occurred: Error en validObject(.Object) : invalid class "SpatialPoints" object: bbox should never contain infinite valuesCalls: source ... SpatialPoints -> new -> initialize -> initialize -> validObject

Ferramenta(s): Representa

Solução(ões): Revise o arquivo de texto “process_info.txt” da pasta “Error” da estrutura de pastas e arquivos das ferramentas CAPFITOGEN. Na linha inferior desse arquivo de texto encontrará que pode aparecer um aviso dizendo “ATENCION!!, error al eliminarse todos los registros FE al considerar datos de otros bancos como no faltantes”. Isto indica que a ferramenta ficou sem dados de fontes externas, pois todos os dados fornecidos provêm de “bancos de germoplasma” e ao se indicar que estes não devem se tomar como faltantes, se produz um erro, ao não haver dados por analisar. Elimine a opção de utilizar fontes externas ou permita que Representa tome dados de outros bancos como faltantes (seção 6.3.1.10).

9. Aviso de erro: An error occurred: Error en dist(x[ss[[i]],], method = metric, ...) : no se permiten vectores de longitud negativaCalls: source ... withVisible -> eval -> eval -> pamk -> distcritmulti -> dist

Ferramenta (s): ELC mapas

Solução(ões): Aparece este erro quando o país ou região é muito grande, a resolução é alta (um tamanho de célula menor) e se pede à ferramenta realizar a determinação do número ótimo de agrupamentos “medoides”. A primeira solução aplicável é executar de novo a análise mas usando o método “elbow”. Se apesar disso, se gera outro erro, a recomendação é usar uma resolução menor (tamanho de célula maior).

10. Aviso de erro: An error occurred: Error en merge.data.frame(as.data.frame(x), as.data.frame(y), ...) : no se permiten vectores de longitud negativaCalls: source ... merge -> merge.default -> merge -> merge.data.frame

Ferramenta (s): ELC mapas

Solução(ões): O erro persiste, pois as matrizes que se geram são tão grandes que o método de determinação ótima do número de agrupamentos “elbow” tão pouco é capaz de manejar esse tamanho de matrizes. A solução já passa por usar uma resolução menor (um tamanho de célula maior).

11. Aviso de erro: An error occurred: Error en .checkNumericCoerce2double(obj) : cannot retrieve coordinates from non-numeric elementsCalls: source ... coordinates -> .local -> do.call -> .checkNumericCoerce2double

Ferramenta: GEOQUAL

Solução(ões): Erro na codificação das coordenadas ou na elaboração da tabela de passaporte. Para o primeiro caso, corrija manualmente no Excel e volte a salvar em texto delimitado por tabulações. Para o segundo caso, a ordem das variáveis não é o indicado e por isso as colunas correspondentes às coordenadas se encontram deslocadas. Restringir estritamente a ordem das variáveis ao formato indicado e não acrescente colunas nem mude sua ordem.

12. Aviso de erro: An error occurred: Error en readChar(con, 5L, useBytes = TRUE) : no se puede abrir la conexiónCalls: source -> withVisible -> eval -> eval -> load -> readChar

Ferramenta: Pode apresentar-se em qualquer ferramenta

Solução(ões): Este erro costuma corresponder à introdução errada de parâmetros. Por exemplo, em ELC mapas, se um país como Cuba se indica à ferramenta uma resolução de célula de 10x10 km, pode aparecer este erro. Também quando se indica um caminho equivocado onde se encontra as ferramentas ou as tabelas de passaporte, etc. Para evitar este problema, revise parâmetro por parâmetro, confirmando se os valores são os corretos.

13. Aviso de erro: An error occurred: Error en apply(x, 2, fun2) : dim(X) must have a positive lengthCalls: source ... extract -> .xyValues -> .xyvBuf -> lapply -> FUN -> apply

Ferramenta: Pode apresentar-se em ferramentas com função de extração radial

Solução(ões): Este erro pode dar-se quando o usuário solicita uma extração radial utilizando um raio muito pequeno (parâmetro tamp) com relação ao tamanho de célula ou resolução das variáveis ecogeográficas (parâmetro “resol1”). Por exemplo, solicitar uma extração radial de 1000 m usando resoluções de célula de 10x10 km aprox (5 arc-min). Uma situação assim produz valores nulos de extração que logo geram este erro. Prove com raios de tamanho maior, especialmente maiores ao tamanho de lado de cada célula e/ou tente com uma resolução maior (por exemplo, para uma extração radial de 1000 m passar de “Células 5x5 km aprox (2.5 arc-min)” a “Células 1x1 km aprox (30 arc-seg)” poderia solucionar o problema). Se ainda assim o problema não se solucionar, seria recomendável usar extrações pontuais.

14. Aviso de erro: An error occurred: Error en `colnames<-`(*tmp*, value = "ACCENUMB") : el atributo 'names' [1] debe tener la misma longitud que el vector [0]Calls: source -> withVisible -> eval -> eval -> colnames<-

Ferramenta: Pode apresentar-se em ferramentas onde o usuário ingresse dados de passaporte

Solução(ões): Este aviso de erro pode se produzir quando o usuário indica uma tabela de passaporte no parâmetro “passaporte” que não tenha o número de colunas esperado, devido a que se tenha eliminado acidentalmente alguma coluna, ou que a ferramenta espera colunas adicionais não incluídas (como é o caso de ColNucleo que espera a coluna adicional “AVAILAB”) ou que se indica no parâmetro “geoqual” que a tabela contém as quatro colunas adicionais produto da análise da ferramenta GEOQUAL e na realidade não as possui. Certificar-se do conteúdo da tabela passaporte que ingresa e de acordo a isso, usar o parâmetro “geoqual”.

15. Aviso de erro: An error occurred: Error en if (any(puntosorig\$DECLATITUDE >= 90 | puntosorig\$DECLATITUDE <= : valor ausente donde TRUE/FALSE es necesarioCalls: source -> withVisible -> eval -> eval ó
An error occurred: Error en if (any(puntosorig\$DECLONGITUDE >= 180 | puntosorig\$DECLONGITUDE <= : valor ausente donde TRUE/FALSE es necesarioCalls: source -> withVisible -> eval -> eval

Ferramenta: Pode apresentar-se em ferramentas onde o usuário ingresse dados de passaporte

Solução(ões): Existe um erro nas coordenadas de ao menos uma entrada, e isso pode ser devido aos próprios erros da codificação das coordenadas ou a que inclusive não tenha valor nenhum (incluindo NA) nos campos das coordenadas. Para solucionar o problema no primeiro caso, verifique a codificação das coordenadas em sexagesimal, que estejam de acordo com o formato FAO/Bioversity 2012 e para decimal, que os valores para DECLATITUDE estejam entre -90 e 90 e para DECLONGITUDE entre -180 e 180. Se o erro se deve ao segundo caso (campos vazios), isto pode ser devido à aparição de entradas “fantasma”, as quais se formam ao se criar a tabela de passaporte no excel e haver algumas filas a mais, que lamentavelmente não podem ser facilmente identificadas porque aparecem em branco, mas que ao se exportar a tabela em formato de texto, nesse momento aparecem. O sistema entende que são entradas, porque ocupam uma fila, mas não têm dados em absoluto, incluindo coordenadas, por isso se gera este erro.

11. Créditos

11.1 Como citar CAPFITOGEN

A citação das ferramentas CAPFITOGEN ou do presente manual de usuário deve ser feita como se segue:

Parra-Quijano, M., Torres, E., Iriondo, J.M., López, F. 2014. Manual de usuário ferramentas CAPFITOGEN versão 1.2. Tratado Internacional sobre os Recursos Fitogenéticos para a Alimentação e a Agricultura, FAO, Roma.

11.2 Programas informáticos usados em CAPFITOGEN

O desenvolvimento das ferramentas CAPFITOGEN tem sido possível graças ao financiamento do Ministerio de Asuntos Exteriores y de Cooperación de España e o Tratado Internacional sobre os Recursos Fitogenéticos para a Alimentação e a Agricultura.

As ferramentas CAPFITOGEN estão suportadas por R versão 2.15.2 (<http://cran.r-project.org/>).

R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>

Para a interface facilitada para usuários de GEOQUAL se utilizou o Rwui (<http://sysbio.mrc-bsu.cam.ac.uk/Rwui/>).

O desenvolvimento do instalador e a melhora da interface de visualização foram realizados por Lextrend S.L. (<http://www.lextrend.com/>).

11.3 Pacotes R

As ferramentas CAPFITOGEN usam os seguintes pacotes R:

sp (Edzer Pebesma, Roger Bivand, Barry Rowlingson, Virgilio Gomez-Rubio)

raster (Robert J. Hijmans e Jacob van Etten)

maptools (Roger Bivand, Nicholas Lewin-Koh)

rgdal (Roger Bivand, Keitt Tim, Rowlingson Barry)

rgeos (Roger Bivand, Colin Rundel, Edzer Pebesma, Karl Ove Hufthammer)

RJSONIO (Duncan Temple Lang)

googleVis (Markus Gesmann, Diego de Castillo)

cluster (Martin Maechler)

modeltools (Torsten Hothorn, Friedrich Leisch, Achim Zeileis)

fpc (Christian Hennig)

dismo (Robert J. Hijmans, Steven Phillips, John Leathwick e Jane Elith)

ade4 (Daniel Chessel, Anne-Beatrice Dufour e Stephane Dray)

labdsv (David W. Roberts)

vegan (Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens e Helene Wagner)

Estes pacotes por sua vez dependem de outros pacotes para seu correto funcionamento. Agradecemos a colaboração de todos os autores relacionados.

Alguns mapas de diversidade genotípica usam parâmetros derivados dos scripts AFLPdat (<http://www.nhm.uio.no/english/research/ncb/aflpdat/>) de Dorothée Ehrich.

11.4 Dados

A fonte de informação administrativa de alta precisão provém de:

Global Administrative Areas GADM versão 2 (<http://www.gadm.org>)

Os mapas de anéis de 1 e 10 km ao redor de áreas administrativas de alta precisão (GADM) se obtiveram através da função Buffer (zona de influência) em ArcGIS 10.

O mapa de anéis de 1, 10 e 20 km de baixa precisão são uma modificação do mapa ESRI world countries (2011) que se encontra licenciado pela Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License.

O mapa de uso do solo utilizado para o cálculo do parâmetro SUITQUAL é

Global Land Cover (GLC) 2000 e sua referência é:

Bartholome E., Belward A.S., Achard F., Bartalev S., Carmona Moreno C., Eva H., Fritz S., Grégoire J.-M., Mayaux P. & Stibig H.-J. (2002). Global Land Cover mapping for the year 2000

– Project status November 2002, Office for Official Publications of the European Communities, Luxembourg EUR 20524).

A fonte de informação (variáveis) ecogeográfica provém de:

Worldclim (<http://www.worldclim.org>)

Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G. and Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25:1965-1978.

Harmonized world soil database

(<http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/>)

FAO/IIASA/ISRIC/ISSCAS/JRC. 2012. Harmonized World Soil Database (version 1.2). FAO, Rome, Italy and IIASA, Laxenburg, Austria.

Modelos digitais de elevação (MDE) do Shuttle Radar Topography Mission (SRTM)

(<http://srtm.csi.cgiar.org/>)

Jarvis, A., H.I. Reuter, A. Nelson, E. Guevara, 2008, Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database:<http://srtm.csi.cgiar.org>.

11.5 Metodologias

A avaliação da qualidade da georreferenciação em dados de passaporte é uma técnica originalmente desenvolvida no projeto SIERFE (Sistema de Informação Ecogeográfica para os Recursos Fitogenéticos Espanhóis). O desenvolvimento de GEOQUAL, os mapas de caracterização ecogeográfica do território e o conceito de Representatividade Ecogeográfica (RE) são um aporte original de Mauricio Parra Quijano, Elena Torres Lamas e José María Iriondo Alegría.

O conceito original considerado para o desenvolvimento de mapas de diversidade (ferramentas DIV mapas) foi publicado por:

van Zonneveld M, Scheldeman X, Escribano P, Viruel MA, Van Damme P, et al. (2012) Mapping Genetic Diversity of Cherimoya (*Annona cherimola* Mill.): Application of Spatial Analysis for Conservation and Use of Plant Genetic Resources. *PLoS ONE* 7(1): e29845. doi:10.1371/journal.pone.0029845

A ferramenta FIGS_R incorpora ideias e desenvolvimentos alcançados em PGR Secure (<http://www.pgrsecure.org>), um projeto colaborativo financiado pelo Sétimo Programa Marco (THEME KBBE 2010.1.1-03, “Characterization of biodiversity resources for wild crop relatives to improve crops by breeding). Os conceitos e desenvolvimentos introduzidos em FIGS_R provêm do trabalho do grupo de

“Predictive characterization” (Task 2.2) del WP2 “Informatics”, liderado por Bioversity International (<http://www.bioversityinternational.org/>). Os pesquisadores envolvidos nestes desenvolvimentos são: Imke Thormann, Jacob van Etten e Sonia Dias (Bioversity), José Iriando e Luisa Rubio (Universidad Rey Juan Carlos), Shelagh Kell (University of Birmingham), Dag Endresen (GBIF), Rosa García (CRF-INIA) e Mauricio Parra Quijano (TIRFAA).

11.6 Outros agradecimentos

Agradecimentos a Fernando Latorre (CRF-INIA, España) por seu firme apoio ao Programa CAPFITOGEN e ao desenvolvimento das ferramentas. Agradecimentos, também por apoio e comentários, a outros pesquisadores do CRF-INIA, em particular Lucía de la Rosa, Rosa García e Luis Ayerbe.

Agradecimentos a Robert J. Hijmans, por sua valiosa colaboração na resolução de dúvidas com certos pacotes de R e sua generosidade ao permitir distribuir a informação de worldclim e GADM dentro das ferramentas CAPFITOGEN.

Também agradecemos a múltiplos centros de pesquisa e conservação de recursos genéticos vegetais e a muitos de seus pesquisadores e curadores por seus valiosos comentários e contribuições para a melhora das ferramentas CAPFITOGEN.

12. Anexos

12.1 Variáveis ecogeográficas disponíveis. Variáveis bioclimáticas.

Código	Descrição de variáveis	Unidade	Fonte
prec_1	Precipitação média de janeiro	mm	Worldclim
prec_2	Precipitação média de fevereiro	mm	Worldclim
prec_3	Precipitação média de março	mm	Worldclim
prec_4	Precipitação média de abril	mm	Worldclim
prec_5	Precipitação média de maio	mm	Worldclim
prec_6	Precipitação média de junho	mm	Worldclim
prec_7	Precipitação média de julho	mm	Worldclim
prec_8	Precipitação média de agosto	mm	Worldclim
prec_9	Precipitação média de setembro	mm	Worldclim
prec_10	Precipitação média de outubro	mm	Worldclim
prec_11	Precipitação média de novembro	mm	Worldclim
prec_12	Precipitação média de dezembro	mm	Worldclim
tmean_1	Temperatura média de janeiro	°C	Worldclim
tmean_2	Temperatura média de fevereiro	°C	Worldclim
tmean_3	Temperatura média de março	°C	Worldclim
tmean_4	Temperatura média de abril	°C	Worldclim
tmean_5	Temperatura média de maio	°C	Worldclim
tmean_6	Temperatura média de junho	°C	Worldclim
tmean_7	Temperatura média de julho	°C	Worldclim
tmean_8	Temperatura média de agosto	°C	Worldclim
tmean_9	Temperatura média de setembro	°C	Worldclim
tmean_10	Temperatura média de outubro	°C	Worldclim
tmean_11	Temperatura média de novembro	°C	Worldclim
tmean_12	Temperatura média de dezembro	°C	Worldclim
tmin_1	Temperatura mínima de janeiro	°C	Worldclim
tmin_2	Temperatura mínima de fevereiro	°C	Worldclim
tmin_3	Temperatura mínima de março	°C	Worldclim
tmin_4	Temperatura mínima de abril	°C	Worldclim
tmin_5	Temperatura mínima de maio	°C	Worldclim
tmin_6	Temperatura mínima de junho	°C	Worldclim
tmin_7	Temperatura mínima de julho	°C	Worldclim
tmin_8	Temperatura mínima de agosto	°C	Worldclim
tmin_9	Temperatura mínima de setembro	°C	Worldclim
tmin_10	Temperatura mínima de outubro	°C	Worldclim
tmin_11	Temperatura mínima de novembro	°C	Worldclim
tmin_12	Temperatura mínima de dezembro	°C	Worldclim
tmax_1	Temperatura máxima de janeiro	°C	Worldclim
tmax_2	Temperatura máxima de fevereiro	°C	Worldclim
tmax_3	Temperatura máxima de março	°C	Worldclim
tmax_4	Temperatura máxima de abril	°C	Worldclim
tmax_5	Temperatura máxima de maio	°C	Worldclim
tmax_6	Temperatura máxima de junho	°C	Worldclim
tmax_7	Temperatura máxima de julho	°C	Worldclim
tmax_8	Temperatura máxima de agosto	°C	Worldclim
tmax_9	Temperatura máxima de setembro	°C	Worldclim
tmax_10	Temperatura máxima de outubro	°C	Worldclim
tmax_11	Temperatura máxima de novembro	°C	Worldclim

12.1 Continuação

Código	Descrição de variáveis	Unidade	Fonte
tmax_12	Temperatura máxima de dezembro	°C	Worldclim
bio_1	Temperatura média anual	°C	Worldclim
bio_2	Categoria média de temperaturas diurnas	°C	Worldclim
bio_3	Isotermalidade (BIOCLIM2/BIOCLIM7)(*100)		Worldclim
bio_4	Sazonalidade na temperatura (desvio padrão*100)		Worldclim
bio_5	Máxima temperatura do mês mais quente	°C	Worldclim
bio_6	Mínima temperatura do mês mais frio	°C	Worldclim
bio_7	Categoria de temperatura anual (BIOCLIM5 - BIOCLIM6)	°C	Worldclim
bio_8	Temperatura média do trimestre mais úmido (3 meses mais chuvosos)	°C	Worldclim
bio_9	Temperatura média do trimestre mais seco (3 meses mais secos)	°C	Worldclim
bio_10	Temperatura média do mês mais quente (3 meses mais quentes)	°C	Worldclim
bio_11	Temperatura média do trimestre mais frio (3 meses mais frios)	°C	Worldclim
bio_12	Precipitação anual	mm	Worldclim
bio_13	Precipitação do mês mais úmido	mm	Worldclim
bio_14	Precipitação do mês mais seco	mm	Worldclim
bio_15	Sazonalidade na precipitação (coeficiente de variação)	mm	Worldclim
bio_16	Precipitação do trimestre mais úmido (3 meses mais chuvosos)	mm	Worldclim
bio_17	Precipitação do trimestre mais seco (3 meses mais secos)	mm	Worldclim
bio_18	Precipitação do trimestre mais quente (3 meses mais quentes)	mm	Worldclim
bio_19	Precipitação do trimestre mais frio (3 meses mais frios)	mm	Worldclim

12.2 Variáveis ecogeográficas disponíveis. Variáveis edáficas.

Código	Descrição de variáveis	Unidade	Fonte
ref_depth	Profundidade referência da unidade de solo	m	HWS Database
t_gravel	Conteúdo de cascalho em solo superficial	%vol.	HWS Database
t_sand	Conteúdo de areia em solo superficial	% peso	HWS Database
t_silt	Conteúdo de limo em solo superficial	% peso	HWS Database
t_clay	Conteúdo de argila em solo superficial	% peso	HWS Database
t_ref_bulk	Densidade aparente de referência em solo superficial	kg/dm3	HWS Database
t_oc	Conteúdo de carvão orgânico em solo superficial	% peso	HWS Database
t_ph_h2o	pH em solo superficial em solução solo-água	-log(H+)	HWS Database
t_cec_clay	Capacidade de intercâmbio catiónico de argila em solo superficial	cmol/kg	HWS Database
t_cec_soil	Capacidade de intercâmbio catiónico em solo superficial (geral)	cmol/kg	HWS Database
t_bs	Saturação de bases em solo superficial.	%	HWS Database
t_teb	Bases intercambiáveis totais em solo superficial	cmol/kg	HWS Database
t_caco3	Carbonato de cálcio em solo superficial	% peso	HWS Database
t_caso4	Gessos em solo superficial	% peso	HWS Database
t_esp	Sodicidade em solo superficial	%	HWS Database
t_ece	Salinidade em solo superficial	dS/m	HWS Database
s_gravel	Conteúdo de cascalho em subsolo	%vol	HWS Database
s_sand	Conteúdo de areia em subsolo	% peso	HWS Database
s_silt	Conteúdo de limo em subsolo	% peso	HWS Database
s_clay	Conteúdo de argila em subsolo	% peso	HWS Database
s_ref_bulk	Densidade aparente de referência em subsolo	kg/dm3	HWS Database
s_oc	Conteúdo de carvão orgânico em subsolo	% peso	HWS Database
s_ph_h2o	pH em subsolo em solução solo-água	-log(H+)	HWS Database
s_cec_clay	Capacidade de intercâmbio catiónico de argila em subsolo	cmol/kg	HWS Database

12.2 Continuação

Código	Descrição de variáveis	Unidade	Fonte
s_cec_soil	Capacidade de intercâmbio catiónico em subsolo (geral)	cmol/kg	HWS Database
s_bs	Saturação de bases em subsolo.	%	HWS Database
s_teb	Bases intercambiáveis totais em subsolo	cmol/kg	HWS Database
s_caco3	Carbonato de cálcio em subsolo	% peso	HWS Database
s_caso4	Gessos em subsolo	% peso	HWS Database
s_esp	Sodicidade em subsolo	%	HWS Database
s_ece	Salinidade em subsolo	dS/m	HWS Database

12.3 Variáveis ecogeográficas disponíveis. Variáveis geofísicas.

Código	Descrição de variáveis	Unidade	Fonte
alt	Elevação. Metros sobre o nível do mar	m	Worldclim
slope	Declive (em graus) da superfície do terreno	°	SRTM MDE
aspect	Orientação (em graus) da superfície do terreno	°	SRTM MDE
northness	Norticidade. 1 se a orientação tende ao norte, -1 ao sul		SRTM MDE
eastness	Esticidad. 1 se a orientação tende a leste, -1 a oeste		SRTM MDE
POINT_X	Longitude	°	
POINT_Y	Latitude	°	

Nota: As direções web das fontes de informação (worldclim, SRTM MDE e HWS Database) aparecem no capítulo 11 (Créditos).

12.4 Explicação das colunas adicionais na tabela de resultados “tabla_de_analisisGEOQUAL.txt”.

Variável	Explicação
globlandc	Valor extraído de GLC 2000(Global Land Cover 2000).
DISTOLAND	Anel de distância no qual caem as coordenadas. (0= terra, 1=1km, 10=10km, etc.).
SUITQUAL	Parâmetro SUITQUAL (valores de 0 a 20).
ID_0	Valor extraído de GADM que identifica o polígono de país.
ISO	Valor extraído de GADM que se compara com ORIGCTY.
NAME_0	Valor extraído de GADM que corresponde à denominação completa de país.
ID_1	Valor extraído de GADM que identifica ao polígono do nível NAME_1.
NAME_1	Valor extraído de GADM que se compara com ADM1.
VARNAME_1	Valor extraído de GADM para nomes alternativos a NAME_1.
ENGTYPE_1	Valor extraído de GADM que define o tipo de administração que representa NAME_1 .
ID_2	Valor extraído de GADM que identifica ao polígono do nível NAME_2.
NAME_2	Valor extraído de GADM que se compara com ADM2.
VARNAME_2	Valor extraído de GADM para nomes alternativos a NAME_2.
ENGTYPE_2	Valor extraído de GADM que define o tipo de administração que representa NAME_2 .
ID_3	Valor extraído de GADM que identifica ao polígono do nível NAME_3.
NAME_3	Valor extraído de GADM que se compara com ADM3.
VARNAME_3	Valor extraído de GADM para nomes alternativos a NAME_3.
ENGTYPE_3	Valor extraído de GADM que define o tipo de administração que representa NAME_3.
ID_4	Valor extraído de GADM que identifica ao polígono do nível NAME_4.
NAME_4	Valor extraído de GADM que se compara com ADM4.
VARNAME_4	Valor extraído de GADM para nomes alternativos a NAME_4.
ENGTYPE4	Valor extraído de GADM que define o tipo de administração que representa NAME_4.
NIVELMAX	Segundo o país, este é o nível administrativo de menor nível incluído em GADM.
LOCALQUAL	Parâmetro LOCALQUAL (valores de 0 a 20).
COORQUAL	Parâmetro COORQUAL (valores de 0 a 20).
intertemp	Subparâmetro intertemp de COORQUAL.
errores	Subparâmetro errores de COORQUAL.
precis	Subparâmetro precis de COORQUAL.
georble	Subparâmetro georble de COORQUAL.
TOTALQUAL	Parâmetro TOTALQUAL (valores de 0 a 40 ou 0 a 60, dependendo se se inclui ou não LOCALQUAL).
TOTALQUAL100	Parâmetro TOTALQUAL100 (valores de 0 a 100).

12.5 Tabela de disponibilidade de tamanhos de célula por região/país

No momento, existe disponibilidade de informação ecogeográfica adaptada às ferramentas CAPFITOGEN para 164 países, 3 regiões e nível global. Para os países o tamanho de célula disponível é de 30 arc-segundos (~1x1 km em Equador) e 2.5 arc-minutos (~5x5 km em Equador). No caso de alguns países de grande tamanho (como por exemplo Brasil), a ferramenta ELC mapas pode gerar erros ao usar informação de alta resolução (1x1 km), enquanto que para as demais ferramentas não seria um problema seu uso.

No caso das regiões, se dispõe de informação ecogeográfica para Europa a 2.5 arc-minutos (~5x5 km em Equador) e para Mesoamérica, América do sul, e o mundo em duas resoluções, 5 arc-minutos (~10x10 km) e 10 arc-minutos (~20x20 km). Da mesma forma que no caso de países de dimensões maiores a um milhão de quilômetros quadrados, na ferramenta ELC mapas podem aparecer problemas ao usar-se uma resolução de 10x10 km.

Também conta-se com informação ecogeográfica do Brasil por estados a resoluções 1x1 e 5x5 km.

No link seguinte você pode ver uma tabela com as resoluções disponíveis para os países e regiões: <http://www.agrobiodiversidad.org/blog/?p=1091>

GEOQUAL

ELC mapas

ECOGEO

Representa

DIV mapas

ColNucleo

FIGS_R

Com financiamento e apoio de:



**Tratado Internacional sobre os Recursos Fitogenéticos
para a Alimentação e a Agricultura**

Telefone : (+39) 06 5705 6343 • Fax (+39) 06 570 56347 • capfitogen@fao.org

<http://www.planttteaty.org/capfitogen>

ISBN 978-92-5-308255-1



9 789253 082551

I3700S/1/03.14